

Mouse Dnmt3a DNA sequence

1 GAATTCCGGC CTGCTGCCGG GCCGCCGCAC CCGCCGGGCC ACACGGCAGA
51 GCCGCCTGAA GCCCAGCGCT GAGGCTGCAC TTTCCGAGG GCTTGACATC
101 AGGGTCTATG TTTAAGTCTT AGCTCTTGT TACAAAGACC ACGGCAATT
151 CTTCTCTGAA GCCCTCGCAG CCCCACAGCG CCCTCGCAGC CCCAGCCTGC
201 CGCCTACTGC CCAGCAATGC CCTCCAGCGG CCCCCGGGAC ACCAGCAGCT
251 CCTCTCTGGA GCGGGAGGAT GATCGAAAGG AAGGAGAGGA ACAGGAGGAG
301 AACCGTGCA AGGAAGAGCG CCAGGAGCCC AGCGCCACGG CCCGGAAAGGT
351 GGGGAGCCCT GGCCGGAAGC GCAAGCACCC ACCGGTGGAA ACCAGTGACA
401 CCCCCAAGGA CCCACCGAGT ACCACCAAGT CTCAGCCCAT GCCCCAGGAC
451 TCTGGCCCT CAGATCTGCT ACCCAATGGA GACTTGGAGA AGCGGAGTGA
501 ACCCCAACCT GAGGAGGGGA GCCCAGCTGC AGGGCAGAAC GGTGGGGCCC
551 CAGCTGAAGG AGAGGAACT GAGACCCAC CAGAAGCCTC CAGAGCTGTG
601 GAGAATGGCT GCTGTGTGAC CAAGGAAGGC CGTGGAGCCT CTGCAGGAGA
651 GGGCAAAGAA CAGAACAGA CCAACATCGA ATCCATGAAA ATGGAGGGCT
701 CCCCCGGCCG ACTGCGAGGT GGCTTGGCT GGGAGTCCAG CCTCCGTCA
751 CGACCCATGC CAAGACTCAC CTTCCAGGCA GGGGACCCCT ACTACATCAG
801 CAAACGGAAA CGGGATGAGT GGCTGGCACG TTGGAAAAGG GAGGCTGAGA
851 AGAAAGCCAA GTAAATTGCA GTAATGAATG CTGTGGAAGA GAACCAGGCC
901 TCTGGAGAGT CTCAGAAGGT GGAGGAGGCC AGCCCTCCTG CTGTGCAGCA
951 GCCCACGGAC CCTGCTTCTC CGACTGTGGC CACCAACCCCT GAGCCAGTAG
1001 GAGGGGATGC TGGGGACAAG AATGCTACCA AAGCAGCCGA CGATGAGCCT
1051 GAGTATGAGG ATGCCGGGG CTTGGCATT GGAGAGCTGG TGTGGGGAA
1101 ACTTCGGGGC TTCTCCTGGT GGCCAGGCCG AATTGTGTCT TGGTGGATGA

FIG. 1A-1

1151 CAGGCCGGAG CCGAGCAGCT GAAGGCACTC GCTGGTCA T GTGGTTGG
1201 GATGGCAAGT TCTCAGTGGT GTGTGTGGAG AACCTCATGC CGCTGACCTC
1251 CTTCTGCAGT GCATTCCACC AGGCCACCTA CAACAAGCAG CCCATGTACC
1301 GCAAAGCCAT CTACGAAGTC CTCCAGGTGG CCAGCAGCCG TCCCAGGAAG
1351 CTGTTCCAG CTTGCCATGA CAGTGATGAA AGTGACAGTG GCAAGGCTGT
1401 GGAAGTGCAG ACAAGCAGA TGATTGAATG GGCCCTCGGT GGCTTCCAGC
1451 CCTCGGGTCC TAAGGGCCTG GAGCCACCAAG AAGAAGAGAA GAATCCTTAC
1501 AAGGAAGTTT ACACCGACAT GTGGGTGGAG CCTGAAGCAG CTGCTTACGC
1551 CCCACCCCCA CCAGCCAAGA AACCCAGAAA GAGCACAAACA GAGAAACCTA
1601 AGGTCAAGGA GATCATTGAT GAGGCCACAA GGGAGCGGCT CGTGTATGAG
1651 GTGCGCCAGA AGTGCAGAAA CATCGAGGAC ATTTGTATCT CATGTGGAG
1701 CCTCAATGTC ACCCTGGAGC ACCCACTCTT CATTGGAGGC ATGTGCCAGA
1751 ACTGTAAGAA CTGCTTCTTG GAGTGTCTT ACCACTATGA CGACGATGGG
1801 TACCAAGTCCT ATTGCACCAT CTGCTGTGGG GGGCGTGAAG TGCTCATGTG
1851 TGGGAACAAAC AACTGCTGCA GGTGCTTTG TGTGAGTGT GTGGATCTCT
1901 TGGTGGGCC AGGAGCTGCT CAGGCACCCA TTAAGGAAGA CCCCTGGAAC
1951 TGCTACATGT GCGGGCATAA GGGCACCTAT GGGCTGCTGC GAAGACGGGA
2001 AGACTGGCCT TCTCGACTCC AGATGTTCTT TGCCAATAAC CATGACCAGG
2051 AATTGACCC CCCAAAGGTT TACCCACCTG TGCCAGCTGA GAAGAGGAAG
2101 CCCATCCCGT TGCTGTCTCT CTTTGATGGG ATTGCTACAG GGCTCCTGGT
2151 GCTGAAGGAC CTGGGCATCC AAGTGGACCG CTACATTGCC TCCGAGGTGT
2201 GTGAGGACTC CATCACGGTG GGCATGGTCC GGCACCAAGGG AAAGATCATG
2251 TACGTGGGG ACGTCCGCAG CGTCACACAG AAGCATATCC AGGAGTGGGG
2301 CCCATTGAC CTGGTATTG GAGGCAGTCC CTGCAATGAC CTCTCCATTG

FIG. 1A-2

2351 TCAACCCTGC CCGCAAGGGA CTTTATGAGG GTACTGGCCG CCTCTTCTT
2401 GAGTTCTACC GCCTCCTGCA TGATGCCGGG CCCAAGGAGG GAGATGATCG
2451 CCCCTTCTTC TGGCTCTTG AGAATGTGGT GGCCATGGC GTTAGTCACA
2501 AGAGGGACAT CTCCGGATT CTTGAGTCTA ACCCCGTGAT GATTGACGCC
2551 AAAGAAAGTGT CTGCTGCACA CAGGGCCCGT TACTTCTGGG GTAACCTCC
2601 TGGCATGAAC AGGCCTTGG CATCCACTGT GAATGATAAC CTGGAGCTGC
2651 AAGAGTGTCT GGAGCACGGC AGAATAGCCA AGTCAGCAA AGTGAGGACC
2701 ATTACCACCA GGTCAAACTC TATAAAGCAG GGCAAAGACC AGCATTCCC
2751 CGTCTTCATG AACGAGAAGG AGGACATCCT GTGGTGCACT GAAATGGAAA
2801 GGGTGTGGG CTTCCCCGTC CACTACACAG ACGTCTCCAA CATGAGCCCG
2851 TTGGCGAGGC AGAGACTGCT GGGCCGATCG TGGAGCGTGC CGGTCACTCC
2901 CCACCTCTTC GCTCCGCTGA AGGAATATT TGCTTGTGT TAAGGGACAT
2951 GGGGGCAAAC TGAAGTAGTG ATGATAAAAA AGTTAAACAA ACAAACAAAC
3001 AAAAAACAAA ACAAAACAAT AAAACACCAA GAACGAGAGG ACGGAGAAAA
3051 GTTCAGCACC CAGAAGAGAA AAAGGAATT AAAGCAAACC ACAGAGGAGG
3101 AAAACGCCGG AGGGCTTGGC CTTGCAAAG GGTTGGACAT CATCTCCTGA
3151 GTTTCAATG TTAACCTTCA GTCCTATCTA AAAAGCAAAA TAGGCCCTC
3201 CCCTTCTTCC CCTCCGGTCC TAGGAGGCCA ACTTTTGTG TTCTACTCTT
3251 TTTCAGAGGG GTTTCTGTT TGTTGGGTT TTTGTTCTT GCTGTGACTG
3301 AAACAAGAGA GTTATTGCAG CAAATCACT AACAAACAAA ACTAGAAATG
3351 CCTTGGAGAG GAAAGGGAGA GAGGGAAAAT TCTATAAAA CTTAAATAT
3401 TGGTTTTTT TTTTTTCTT TTTCTATATA TCTCTTGGT TGTCTCTAGC
3451 CTGATCAGAT AGGAGCACAA ACAGGAAGAG AATAGAGACC CTGGAGGCCA
3501 GAGTCTCCTC TCCCACCCCC CGAGCAGTCT CAACAGCACC ATTCCGGTC

FIG. 1A-3

3551 ATGCAAAACA GAACCCAATC AGCAGCAGGG CGCTGAGAGA ACACCACACC
3601 AGACACTTTC TACAGTATTTC CAGGTGCCTA CCACACAGGA AACCTTGAAG
3651 AAAACCAGTT TCTAGAAGCC GCTGTTACCT CTTGTTACA GTTTATATAT
3701 ATATGATAGA TATGAGATAT ATATATATAA AAGGTACTGT TAACTACTGT
3751 ACATCCCCAC TTCTATAATGG TGCTTTCAAACACAGCGAGAT GAGCAAAGAC
3801 ATCAGCTTCC GCCTGGCCCT CTGTGCAAAG GGTTTCAGCC CAGGATGGGG
3851 AGAGGGGAGC AGCTGGAGGG GGTTTAACA AACTGAAGGA TGACCCATAT
3901 CACCCCCCCAC CCCTGGCCCCA TGCCTAGCTT CACCTGCCAA AAAGGGGCTC
3951 AGCTGAGGTG GTCGGACCCCT GGGGAAGCTG AGTGTGGAAT TTATCCAGAC
4001 TCGCGTGCAA TAACCTTAGA ATATGAATCT AAAATGACTG CCTCAGAAAA
4051 ATGGCTTGAG AAAACATTGT CCCTGATTTT GAATTCGTCA GCCACGTTGA
4101 AGGCCCCCTG TGGGATCAGA AATATTCCAG AGTGAGGGAA AGTGACCCGC
4151 CATTAACCCC NCCTGGAGCA AATAAAAAAA CATAACAAAT GT

FIG. 1A-4

Mouse Dnmt3b1 DNA Sequence

1 GAATTCCGGG CGCCGGGGTT AAGCGGCCA AGTAAACGTA GCGCAGCGAT
51 CGGCGCCCGA GATTGCCGAA CCCGACACTC CGCGCCGCC GCGGGCCAGG
101 ACCCGCCGGC CGATGCCGGC GCCGCCCTAC AGCCAGCCTC ACGACAGGCC
151 CGCTGAGGCT TGTGCCAGAC CTTGGAAACC TCAGGTATAT ACCTTTCCAG
201 ACGCGGGATC TCCCCTCCCC CATCCATAGT GCCTTGGGAC CAAATCCAGG
251 GCCTTCTTC AGGAAACAAT GAAGGGAGAC ACCAGACATC TGAATGAAGA
301 AGAGGGTGCC AGCGGGTATG AGGAGTGCAT TATCGTTAAT GGGAACTTCA
351 GTGACCACTC CTCAGACACCC AAGGATGCTC CCTCACCCCC AGTCTGGAG
401 GCAATCTGCA CAGAGCCAGT CTGCACACCA GAGACCAGAG GCCGCAGGTC
451 AAGCTCCCGG CTGTCTAAGA GGGAGGTCTC CAGCCTTCTG AATTACACGC
501 AGGACATGAC AGGAGATGCA GACAGAGATG ATGAAGTAGA TGATGGAAAT
551 GGCTCTGATA TTCTAATGCC AAAGCTCACC CGTGAGACCA AGGACACCAAG
601 GACCGCGCTCT GAAAGCCCGG CTGTCCGAAC CCCACATAGC AATGGGACCT
651 CCAGCTTCCA GAGGCAAAGA GCCTCCCCCA GAATCACCCCC AGGTGGCAG
701 GGCGGCCACC ATGTGCAGGA GTACCCCTGTG GACTTCCGG CTACCAGGTC
751 TCGGAGACGT CGAGCATCGT CTTCAGCAAG CACGCCATGG TCATCCCTG
801 CCAGCGTCCA CTTCATGGAA GAACTGACAC CTAAGAGCGT CAGTACCCCCA
851 TCAGTTGACT TGAGCCAGGA TGGAGATCAG GACGGTATGG ATACCACACA
901 GGTGGATGCA GAGAGCAGAG ATGGAGACAG CACAGAGTAT CAGGATGATA
951 AAGAGTTGG AATAGGTGAC CTCGTGTGGG GAAAGATCAA GGGCTTCTCC
1001 TGGTGGCCTG CCATGGTGGT GTCCCTGAAA GCCACCTCCA AGCGACAGGC

FIG. 1B-1

1051 CATGCCCGA ATGGCTGGG TACAGTCGT TGGTATGCC AAGTTTCTG
1101 AGATCTCTGC TGACAAACTG GTGGCTCTGG GGCTGTTCAAG CCAGCACCTT
1151 AATCTGGCTA CCTTCATAA GCTGGTTCT TATAGGAAGG CCATGTACCA
1201 CACTCTGGAG AAAGCCAGGG TTCGAGCTGG CAAGACCTTC TCCAGCAGTC
1251 CTGGAGAGTC ACTGGAGGAC CAGCTGAAGC CCATGCTGGA GTGGGCCAC
1301 GGTGGCTTCA AGCCTACTGG GATCGAGGGC CTCAAACCCA ACAAGAACCA
1351 ACCAGTGTT AATAAGTCGA AGGTGCGTCG TTCAAGACAGT AGGAACCTAG
1401 AACCCAGGAG ACCGGAGAAC AAAAGTCGAA GACCCACAAC CAATGACTCT
1451 GCTGCTTCTG AGTCCCCCCC ACCCAAGCGC CTCAAGACAA ATAGCTATGG
1501 CGGGAAGGAC CGAGGGGAGG ATGAGGAGAG CCGAGAACCG ATGGCTTCTG
1551 AAGTCACCAA CAACAAGGGC AATCTGGAAG ACCGCTGTT GTCTGTGGA
1601 AAGAAGAACCC CTGTGTCCTT CCACCCCCCTC TTTGAGGGTG GGCTCTGTCA
1651 GAGTTGCCGG GATCGCTTCC TAGAGCTCTT CTACATGTAT GATGAGGACG
1701 GCTATCAGTC CTACTGCACC GTGTGCTGTG AGGGCCGTGA ACTGCTGCTG
1751 TGCAGTAACA CAAGCTGCTG CAGATGCTTC TGTGTGGACT GTCTGGAGGT
1801 GCTGGTGGGC GCAGGCACAG CTGAGGATGCC CAAGCTGCAG GAACCTGGA
1851 GCTGCTATAT GTGCCCTCCCT CAGCGCTGCC ATGGGGTCTT CCGACGCAGG
1901 AAAGATTGGA ACATGCGCCT GCAAGACTTC TTCACTACTG ATCCTGACCT
1951 GGAAGAATTG GAGCCACCCA AGTTGTACCC AGCAATTCT GCAGCCAAA
2001 GGAGGCCCAT TAGAGTCCTG TCTCTGTTG ATGGAATTGC AACGGGGTAC
2051 TTGGTGCTCA AGGACTTGGG TATTAAAGTG GAAAAGTACA TTGCCTCCGA
2101 AGTCTGTGCA GAGTCATCG CTGTGGAAC TGTTAAGCAT GAAGGCCAGA
2151 TCAAATATGT CAATGACGTC CGGAAAATCA CCAAGAAAAA TATTGAAGAG
2201 TGGGGCCCGT TCGACTTGGT GATTGGTGGG AGCCCATGCA ATGATCTCTC

FIG. 1B-2

2251 TAACGTCAT CCTGCCCGCA AAGGTTATA TGAGGGCACA GGAAGGCTCT
2301 TCTTCGAGTT TTACCACTTG CTGAATTATA CCCGCCCCAA GGAGGGCGAC
2351 AACCGTCCAT TCTTCTGGAT GTTCGAGAAT GTTGTGGCCA TGAAAGTCAA
2401 TGACAAGAAA GACATCTCAA GATTCTGGC ATGTAACCCA GTGATGATCG
2451 ATGCCATCAA GGTGTCTGCT GCTCACAGGG CCCGGTACTT CTGGGGTAAC
2501 CTACCCGGAA TGAACAGGCC CGTGATGGCT TCAAAGAATG ATAAGCTCGA
2551 GCTGCAGGAC TGCCTGGAGT TCAGTAGGAC AGCAAAGTTA AAGAAAGTCC
2601 AGACAATAAC CACCAAGTCG AACTCCATCA GACAGGGCAA AAACCAGCTT
2651 TTCCCTGTAG TCATGAATGG CAAGGACGAC GTTTGTGGT GCACTGAGCT
2701 CGAAAGGATC TTCGGCTTCC CTGCTCACTA CACGGACGTG TCCAACATGG
2751 GCCGGGGGGC CCGTCAGAAG CTGCTGGCA GGTCTGGAG TGTACCGGTC
2801 ATCAGACACC TGTTGCCCT CTTGAAGGAC TACTTGCCT GTGAATAGTT
2851 CTACCCAGGA CTGGGGAGCT CTCGGTCAGA GCCAGTCCCC AGAGTCACCC
2901 CTCCCTGAAG GCACCTCACC TGTCCCCTTT TTAGCTCACC TGTGTGGGGC
2951 CTCACATCAC TGTACCTCAG CTTCTCCTG CTCAGTGGGA GCAGAGCCTC
3001 CTGGCCCTTG CAGGGGAGCC CCGGTGCTCC CTCCGTGTGC ACAGCTCAGA
3051 CCTGGCTGCT TAGAGTAGCC CGGCATGGTG CTCATGTTCT CTTACCCCTGA
3101 AACTTTAAAA CTTGAAGTAG GTAGTAAGAT GGCTTCTTT TACCCCTCCTG
3151 AGTTTATCAC TCAGAACTGA TGGCTAAGAT ACCAAAAAAA CAAACAAAAA
3201 CAGAAACAAA AAACAAAAAA AAACCTCAAC AGCTCTCTTA GTACTCAGGT
3251 TCATGCTGCA AAATCACTTG AGATTTGTT TTTAAGTAAC CCGTGCTCCA
3301 CATTGCTGG AGGATGCTAT TGTGAATGTG GGCTCAGATG ACCAAGGTCA
3351 AGGGGCCAAA AAAAATTCCC CCTCTCCCCC CAGGAGTATT TGAAGATGAT
3401 GTTTATGTT TAAGTCTTCC TGGCACCTTC CCCTTGCTTT GGTACAAGGG

FIG. 1B-3

3451 CTGAAGTCCT GTGGTCTTG TAGCATTCC CAGGATGATG ATGTCAGCAG
3501 GGATGACATC ACCACCTTA GGGCTTTCC CTGGCAGGGG CCCATGTGGC
3551 TAGTCCTCAC GAAGACTGGA GTAGAATGTT TGGAGCTCAG GAAGGGTGGG
3601 TGGAGTGGCC CTCTTCCAGG TGTGAGGGAT ACCAAGGAGG AAGCTTAGGG
3651 AAATCCATTTC CCCACTCCCT CTTGCCAAAT GAGGGGCCA GTCCCCAACAA
3701 GCTCAGGTCC CCAGAACCCC CTAGTTCTC ATGAGAAGCT AGGACCAGAA
3751 GCACATCGTT CCCCTTATCT GACCACTGTT TGGGAACTA CAGTAAAAC
3801 CTTCTGGAGA TGTAAAAGC TTTTACCCC ACGATAGATT GTGTTTTAA
3851 GGGGTGCTTT TTTAGGGGC ATCACTGGAG ATAAGAAAGC TGCATTTCAG
3901 AAATGCCATC GTAATGGTTT TTAAACACCT TTTACCTAAT TACAGGTGCT
3951 ATTTATAGA AGCAGACAAC ACTTCTTTT ATGACTCTCA GACTTCTATT
4001 TTCATGTTAC CATTTCCTT GTAACTCGCA AGGTGTGGC TTTGTAACT
4051 TCACAGGTGT GGGGAGAGAC TGCCTTGTGTT CAACAGTTG TCTCCACTGG
4101 TTTCTAATT TTAGGTGCAA AGATGACAGA TGCCCAGAGT TTACCTTCT
4151 GGTTGATTAA AGTTGTATT CTCTAAAAAA AAAAAAAA AAAAA

FIG. 1B-4

Human DNMT3A DNA Sequence

1 GCCGGCGG CACCAGGGCG CGCAGCCGGG
28 CCGGCCCCAC CCCACCGGCC ATACGGTGGA CCCATCGAAC CCCCCACCCA
78 CAGGCTGACA GAGGCACCGT TCACCAGAGG GCTAACACCC GGGATCTATG
128 TTTAACTTT AACTCTCGCC TCCAAAGACC ACCATAATT CTTCCCCAAA
178 GCCCAAGCAGC CCCCCAGCCC CGGGCAGCCC CAGCCTGCCT CCCGGCGCCC
228 AGATGCCCGC CATGCCCTCC AGCGGGCCCCG GGGACACCAAG CAGCTCTGCT
278 GCGGAGCCGG AGGAGGACCG AAAGGACGGA GAGGAGCAGG AGGAGCCCG
328 TGGCAAGGAG GAGGCCAAG AGCCCAGCAC CACGGCACGG AAGGTGGGGC
378 GGCCTGGGAG GAAGGCCAAG CACCCCCCGG TGAAAGCCG TGACACGCCA
428 AAGGACCTG CGGTGATCTC CAAGTCCCCA TCCATGGCCC AGGACTCAGG
478 CGCCTCAGAG CTATTACCCA ATGGGGACTT GGAGAAGCCG AGTGAGCCCC
528 AGCCAGAGGA GGGGAGCCCT GCTGGGGGGC AGAAGGGGGG GGCCCCAGCA
578 GAGGGAGAGG GTGCAGCTGA GACCCTGCCT GAAGCCTCAA GACCAGTGG
628 AAATGGCTGC TGCACCCCCA AGGAGGGCCG AGGAGCCCT GCAGAAGCCG
678 GCAAAGAACAA GAAGGAGACC AACATCGAAT CCATGAAAAT GGAGGGCTCC
728 CGGGGCCGGC TCGGGGTGG CTTGGGCTGG GAGTCCAGCC TCCGTCAAGCG
778 GCCCATGCCG AGGCTCACCT TCCAGGGGGG GGACCCCTAC TACATCAGCA
828 ACCCCAAGCC GGACCGAGTGG CTGGCACGCT GGAAAAGGGA GGCTGAGAAG
878 AAAGCCAAGG TCAGTGCAGG AATGAATGCT GTGGAAGAAA ACCAGGGGCC
928 CGGGGAGTCT CAGAAGGTGG AGGAGGCCAG CCCTCCTGCT GTGCAGCAGC
978 CCACTGACCC CGCATCCCCC ACTGTGGCTA CCACGCCCTGA GCCCGTGGGG
1028 TCCGATGCTG GGGACAAGAA TGCCACCAAA GCAGGCCATG ACGAGCCAGA

FIG. 1C-1

1078 GTACGAGGAC GGCCGGGGCT TTGGCATTGG GGAGCTGGTG TGGGGAAAC
1128 TCGGGGCCTT CTCTGGTGG CCAGGCCGA TTGTGTCTTG GTGGATGACG
1178 GGCCGGAGCC GAGCAGCTGA AGGCACCCCGC TGGGTCAATGT GGTTCCGAGA
1228 CGGCAAATTTC TCAGTGGTGT GTGTTGAGAA GCTGATGCCG CTGAGCTCGT
1278 TTTGCAGTGC GTTCCACCAG GCCACGTACA ACAACGAGCC CATGTACCGC
1328 AAAGCCATCT ACGAGGTCTT GCAGGTGCC AGCAGCCGGC CGGGGAAAGCT
1378 GTTCCCGGTG TGCCACCGACA GCGATGAGAG TGACACTGCC AAGGCCGTCC
1428 AGGTGCAGAA CAAGCCCCATG ATTGAATGGG CCCTGGGGGG CTTCCAGCCT
1478 TCTGGCCCTA AGGGCCTGGA GCCACCAGAA GAAGAGAAGA ATCCCTACAA
1528 AGAAAGTGTAC ACGGACATGT GGGTGGAACCC TGAGGCAGCT GCCTACGCAC
1578 CACCTCCACC AGCCAAAAAG CCCCGGAAGA GCACAGCGGA GAAGCCCAAG
1628 GTCAAGGAGA TTATTGATGA CGGCACAAGA GAGCCGCTGG TGTACGAGGT
1678 GCGGCAGAAG TGCCCGAACCA TTGAGGACAT CTGCATCTCC TGTGGGAGCC
1728 TCAATGTTAC CCTGGAACAC CCCCTCTTCG TTGGAGGAAT GTGCCAAAAC
1778 TGCAAGAACT GCTTCTGGA GTGTGGTAC CAGTACGACG ACCACGGCTA
1828 CCAGTCCTAC TGCACCATCT GCTGTGGGGG CCGTGAGGTG CTCATGTGCG
1878 GAAACAACAA CTGCTGCAGG TGCTTTGCG TGGACTGTGT GGACCTCTTG
1928 GTGGGGCCGG GGGCTGCCA GGCAAGCATT AAGGAAGACC CCTGGAACTG
1978 CTACATGTGC GGGCACAAGG GTACCTACGG GCTGCTGCCG CGGGGAGAGG
2028 ACTGGCCCTC CGGGCTCCAG ATGTTCTTCG CTAATAACCA CGACCAGGAA
2078 TTTGACCCCTC CAAAGGTTA CCCACCTGTC CCAGCTGAGA AGAGGAAGCC
2128 CATCCGGGTG CTGTCCTCT TTGATGGAAT CGCTACAGGG CTCCCTGGTGC
2178 TGAAGGACTT GGGCATTCAAG GTGGACCGCT ACATTGCCTC GGAGGTGTG

FIG. 1C-2

2228 GAGGACTCCA TCACGGTGGG CATGGTCCGG CACCAGGGGA AGATCATGTA
2278 CGTCGGGGAC GTCCGCAGCG TCACACAGAA GCATATCCAG GAGTGGGGCC
2328 CATTGATCT GGTGATTGGG GGCAGTCCT GCAATGACCT CTCCATCGTC
2378 AACCTGCTC GCAAGGGCCT CTACGAGGGC ACTGCCGGC TCTTCTTGA
2428 GTTCTACCGC CTCCCTGCATG ATGCCGGGC CAAGGAGGGA GATGATGCC
2478 CCTTCTTCTG GCTCTTGAC AATGTGGTGG CCATGGGCGT TAGTGACAAG
2528 AGGGACATCT CGCGATTCT CGAGTCCAAC CCTGTGATGA TTGATGCCAA
2578 AGAAAGTGTCA GCTGCACACA GGGCCCGCTA CTTCTGGGT AACCTTCCC
2628 GTATGAACAG GCCGTTGGCA TCCACTGTGA ATGATAAGCT GGAGCTGCAG
2678 GAGTGTCTGG ACCATGCCAG GATAGCCAAG TTCAGCAAAG TGAGGACCAT
2728 TACTACGAGG TCAAACCCA TAAAGCAGGG CAAAGACCAAG CATTTCCTG
2778 TCTTCATGAA TGAGAAAGAC GACATCTTAT GGTGCACTGA AATGAAAGG
2828 GTATTTGGTT TCCCAGTCCA CTATACTGAC GTCTCCAACA TGAGCCGCTT
2878 GCCGAGGCAG AGACTGCTGG GCCGGTCATG GAGCGTGCCA GTCATCCGCC
2928 ACCTCTTCGC TCCGCTGAAG GACTATTTG CGTGTGTGA AGGGACATGG
2978 GGGCAAAC TG AGGTAGCGAC ACAAAAGTTAA ACAAAACAAAC AAAAACACA
3028 AAACATAATA AAACACCAAG AACATGAGGA TGGAGAGAAG TATCAGCACC
3078 CAGAAGAGAA AAAGGAATT AAAACAAAAA CCACAGAGGC GGAAATACCG
3128 GAGGGCTTTG CCTTGCAGAA AGGGTTGGAC ATCATCTCCT GATTTTCAA
3178 TGTATTCTT CAGTCCTATT TAAAAACAAA ACCAAGCTCC CTTCCCTTCC
3228 TCCCCCTTCC CTTTTTTTC GGTAGACCT TTTATTTCT ACTCTTTCA
3278 GAGGGGTTT CTGTTGTTT GGGTTTGTT TCTTGTG ACTGAAACAA
3328 GAAGGTTATT GCAGCAAAAA TCAGTAACAA AAAATAGTAA CAATACCTG
3378 CAGAGGAAAG GTGGGAGGAG AGGAAAAAG GGAAATTTT AAAGAAATCT

FIG. 1C-3

3428 ATATATTGGG TTGTTTTTT TTTTGTTTT TGTTTTTTT TTTGGGTTT
3478 TTTTTTTTA CTATATATCT TTTTTTGTT GTCTCTAGCC TGATCAGATA
3528 GGAGCACAAAG CAGGGGACCG AAAGAGAGAG ACACTCAGGC GGCAAGCATTC
3578 CCTCCCAGCC ACTGAGCTGT CGTGCCAGCA CCATTCTGG TCACGCAAAA
3628 CAGAACCCAG TTAGCAGCAG GGAGACGAGA ACACCACACA AGACATTTT
3678 CTACAGTATT TCAGGTGCCT ACCACACAGG AACACCTGAA GAAAATCAGT
3728 TTCTAGAAGC CGCTGTTACC TCTTGTAC ACTTTATATA TATATGATAG
3778 ATATGAGATA TATATATAAA AGGTACTGTT AACTACTGTA CAACCCGACT
3828 TCATAATGGT GCTTCAAAC AGCGAGATGA GTAAAAACAT CAGCTTCCAC
3878 GTTGCCTTCT GCCCAAAGGG TTTCACCAAG GATGGAGAAA GGGAGACAGC
3928 TTGCAGATGG CGCGTTCTCA CGGTGGGCTC TTCCCTTGG TTTGTAACGA
3978 AGTGAAGGAG GAGAACTTGG GAGCCAGGTT CTCCCTGCCA AAAAGGGGGC
4028 TAGATGAGGT GGTGGGGCCC GTGGACAGCT GAGAGTGGGA TTCATCCAGA
4078 CTCATGCAAT AACCCTTGA TTGTTTCTA AAAGGAGACT CCCTCGCAA
4128 GATGGCAGAG GGTACGGAGT CTTCAGGCC AGTTTCTCAC TTTAGCCAAT
4178 TCGAGGGCTC CTTGTGGTGG GATCAGAACT AATCCAGAGT GTGGAAAGT
4228 GACAGTCAAAC ACCCCACCTG GACCAAATAA AAAACATAC AAAACGTA
4278 AAAAAAAAAA AAAAAA

FIG. 1C-4

Human DNMT3B1 DNA Sequence:

1 GGCGGCGAAT TCGGCACGAG CCCTGCACGG CGGCCAGCCG GCCTCCCGCC
51 AGCCAGCCCC GACCCGGCGC TCCGCCGCC AGCCGGGCC CAGCCAGCCC
101 TCGGGCACGA AACCATGAAG GGAGACACCA GGATCTCAA TGGAGAGGAG
151 GACGCCGGCG GGAGGGAAAGA CTGGATCCTC GTCAACGGGG CCTGCAGCGA
201 CCAGTCCTCC GACTCGCCCC CAATCCTGGA GGCTATCCGC ACCCCGGAGA
251 TCAGAGGCCG AAGATCAAGC TCGCGACTCT CCAAGAGGGA GGTGTCCACT
301 CTGCTAACGCT ACACACAGGA CTTGACAGGC GATGGCGACG GGGAAAGATGG
351 GGATGGCTCT GACACCCCAG TCATGCCAAA GCTCTTCCGG GAAACCAGGA
401 CTCGTTCAGA AAGCCCAGCT GTCCGAACTC GAAATAACAA CAGTGTCTCC
451 AGCCGGGAGA CCCACAGGCC TTCCCCACGT TCCACCCGAG GCCGGCAGGG
501 CCGCAACCAT GTGGACGAGT CCCCCGTGGA GTTCCCGGCT ACCAGGTCCC
551 TGAGACGGCG GGCAACAGCA TCGGCAGGAA CGCCATGGCC GTCCCCCTCCC
601 AGCTCTTACC TTACCATCGA CCTCACAGAC GACACAGAGG ACACACATGG
651 GACGCCCCAG ACCAGCAGTA CCCCCTACGC CCGCCTAGCC CAGGACAGCC
701 AGCAGGGGGG CATGGAGTCC CCGCAGGTGG AGGCAGACAG TGGAGATGGA
751 GACAGTTCAAGT ATATCAGGA TGGGAAGGAG TTTGGAATAG GGGACCTCGT
801 GTGGGGAAAG ATCAAGGGCT TCTCCTGGTG GCCCCCCATG GTGGTGTCTT
851 GGAAGGCCAC CTCCAAGCGA CAGGCTATGT CTGGCATGCC GTGGGTCCAG
901 TGGTTGGCG ATGGCAAGTT CTCCGAGGTC TCTGCAGACA AACTGGTGGC
951 ACTGGGGCTG TTCAGCCAGC ACTTTAATTG GGCCACCTTC AATAAGCTCG
1001 TCTCCTATCG AAAAGCCATG TACCATGCTC TGGAGAAAGC TAGGGTGCCA
1051 GCTGGCAAGA CCTTCCCCAG CAGCCCTGGA GACTCATTGG AGGACCAGCT
1101 GAAGCCCCATG TTGGAGTGGG CCCACGGGGG CTTCAAGGCC ACTGGGATCG
1151 AGGGCCTCAA ACCCAACAAC ACGCAACCAG TGGTTAATAA GTCGAAGGTG

FIG. 1D-1

1201 CGTCGTGCAG GCAGTAGGAA ATTAGAATCA AGGAAATACG AGAACAAAGAC
1251 TCGAAGACGC ACAGCTGACG ACTCAGCCAC CTCTGACTAC TGCCCCGCAC
1301 CCAAGCGCCT CAAGACAAAT TGCTATAACA ACGGCAAAGA CCGAGGGGAT
1351 GAAGATCAGA GCCGACAACA AATGGCTTCA GATGTTGCCA ACAACAAGAG
1401 CAGCCTGGAA GATGGCTGTT TGTCTTGTGG CAGGAAAAAC CCCGTGTCC
1451 TCCACCCCTCT CTTGAGGGG GGGCTCTGTC AGACATGCCG GGATCGCTTC
1501 CTTGAGCTGT TTTACATGTA TGATGACCATG CGCTATCAGT CTTACTGCAC
1551 TGTGTGCTGC GAGGGCCGAG AGCTGCTGCT TTGCAGCAAC ACGAGCTGCT
1601 GCCGGTGTTT CTGTGTGGAG TGCCCTGGAGG TGCTGGTGGG CACAGGCACA
1651 GCGGCCGAGG CCAAGCTTCA GGAGCCCTGG AGCTGCTACA TGTGTCTCCC
1701 GCAGCGCTGT CATGGCGTCC TGCGGGGCCG GAAGGACTGG AACGTGGGCC
1751 TGCAGGCCCT CTTCACCACT GACACGGGGC TTGAATACGA AGCCCCAAG
1801 CTGTACCCCTG CCATTCCCGC AGCCCGAAGG CGGCCCATTC GAGTCCTGTC
1851 ATTGTTGAT GGCAATCGCGA CAGGCTACCT AGTCCTCAA GAGTTGGGCA
1901 TAAAGGTAGG AAAGTACGTC GCTTCTGAAG TGTGTGAGGA GTCCATTGCT
1951 GTTGGAACCG TGAACCACGA GGGGAATATC AAATACGTGA ACGACGTGAG
2001 GAACATCACA AAGAAAAATA TTGAAGAATG GGGCCCATTT GACTTGGTGA
2051 TTGGCGGAAG CCCATCCAAC GATCTCTCAA ATGTGAATCC AGCCAGGAAA
2101 GGCCTGTATG AGGGTACAGG CCGGCTCTTC TTCGAATTT ACCACCTGCT
2151 GAATTACTCA CGCCCCAAGG AGGGTGATGA CCGGCCGTTC TTCTGGATGT
2201 TTGAGAATGT TGTAGCCATG AAGGTTGGCG ACAAGAGGGA CATCTCACGG
2251 TTCCCTGGAGT GTAATCCAGT GATGATTGAT GCCATCAAAG TTTCTGCTGC
2301 TCACAGGGCC CGATACTTCT GGGGCAACCT ACCCGGGATG AACAGGCCCG
2351 TGATAGGATC AAAAGAATGAT AAAACTCGAGC TGCAGGACTG CTTGGAATAC
2401 AATAGGATAG CCAAGTTAAA GAAAGTACAG ACAATAACCA CCAAGTCGAA

FIG. 1D-2

2451 CTCGATCAA CAGGGAAAA ACCAACTTT CCCTGTTGTC ATGAATGGCA
2501 AAGAAGATGT TTTGTTGTC ACTGAGCTCG AAAGGATCTT TGGCTTCCT
2551 GTGCACTACA CAGACGTGTC CAACATGGGC CGTGGTCCCC GCCAGAAGCT
2601 GCTGGGAAGG TCCTGGAGCG TCCCTGTCA CCGACACCTC TTGGCCCCCTC
2651 TGAAGGACTA CTTGCATGT GAATAGTTCC AGCCAGGCC CAAGCCCAC
2701 GGGGTGTGTG GCAGAGCCAG GACCCAGGAC GTGTGATTCC TGAAGGCATC
2751 CCCAGGCCCT GCTCTTCCTC AGCTGTGTGG GTCATACCGT GTACCTCAGT
2801 TCCCTCTTGC TCAGTGGGGG CAGAGCCACC TGACTCTTGC AGGGTAGCC
2851 TGAGGTGCCG CCTCCTTGTG CACAAATCAG ACCTGGCTGC TTGGAGCAGC
2901 CTAACACGGT GCTCATTTT TCTTCTCCTA AAACTTAAA ACTTGAAGTA
2951 GGTAGCAACG TGGCTTTTT TTTTCCCTT CCTGGGTCTA CCACTCAGAG
3001 AAACAATGCC TAAGATACCA AAACCACAGT GCCGACAGCT CTCCAATACT
3051 CAGGTTAATG CTGAAAATC ATCCAAGACA GTTATTGCAA GAGTTAATT
3101 TTTGAAAATC GGCTACTGCT ATGTGTTAC AGACGTGTGC AGTTGTAGGC
3151 ATGTAGCTAC AGGACATTT TAAGGGCCCA GGATCGTTT TTCCCAGGGC
3201 AAGCAGAAGA GAAAATGTTG TATATGTCTT TTACCCGGCA CATTCCCTT
3251 GCCTAAATAC AAGGGCTGGA GTCTGCACGG GACCTATTAG AGTATTTCC
3301 ACAATGATGA TGATTCAGC AGGGATGACC TCATCATCAC ATTCAAGGGCT
3351 ATTTTTCCC CCACAAACCC AAGGGCAGGG GCCACTCTTA GCTAAATCCC
3401 TCCCCGTGAC TCCAATAGAA CCCTCTGGG AGCTCAGGAA GGGGTGTGCT
3451 GAGTTCTATA ATATAAGCTG CCATATATT TGTAGACAAG TATGGCTCCT
3501 CCATATCTCC CTCTTCCCTA GGAGAGGAGT GTGAACCAAG GAGCTTAGAT
3551 AAGACACCCCC CTCAAACCCA TTCCCTCTCC AGGAGACCTA CCCTCCACAG
3601 GCACAGGTCC CCAGATGACA AGTCTGCTAC CCTCATTCT CATCTTTTA
3651 CTAAACTCAG AGGCAGTGAC AGCAGTCAGG GACAGACATA CATTCTCAT

FIG. 1D-3

3701 ACCTTCCCCA CATCTGAGAG ATGACAGGGA AACTGCAAA GCTCGGTGCT
3751 CCCTTTGGAG ATTTTTAAT CCTTTTTAT TCCATAAGAA GTCGTTTTA
3801 GGGAGAACGG GAATTCAAGAC AACCTGCATT TCAGAAATGC TGTCTATAATG
3851 GTTTTAACA CCTTTTACTC TTCTTACTGG TGCTATTTG TAGAATAAGG
3901 AACAAACGTTG ACAAGTTTG TGGGGCTTT TATACACTTT TTAAAATCTC
3951 AAACTTCTAT TTTTATGTTT AACGTTTCA TTAAAATTTT TTTGTAACG
4001 GAGCCACGAC GTAACAAATA TGGGGAAAAA ACTGTGCCTT GTTCAACAG
4051 TTTTGCTAA TTTTAGGCT GAAAGATGAC GGATGCCCTAG AGTTTACCTT
4101 ATGTTAATT AAAATCAGTA TTTGTCTAAA AAAAAAAAAA AAAAA

FIG. 1D-4

Mouse Dnmt3a Protein

1 MPSSGPGDTS SSSLEREDDR KEGEEQEENR GKEERQEPSA TARKVGRPGR
51 KRKHPPVESS DTPKDPAVTT KSQPMAQDSG PSDLLPNGDL EKRSEPQPEE
101 GSPAAGQKCGG APAECEGTET PPEASRAVEN GCCVTKEGRG ASAGEGKEQK
151 QTNIIESMKME GSRGRLRGGL GWESSLRQRP MPRLTFQAGD PYYISKRKRD
201 EWLARWKREA EKKAKVIAM NAVEENQASG ESQKVEEASP PAVQQPTDPA
251 SPTVATTPEP VCGDAGDKNA TKAADDEPEY EDGRGFGICE LVWGLRGS
301 WWPGRIVSWW MTGRSRAAEG TRWVMWFGDG KFSVVCVEKL MPLSSFCSAF
351 HQATYNKQPM YRKAIYEVLQ VASSRAGKLF PACHDSDESD SGKAVEVQNK
401 QMIEWALGGF QPSGPKGLEP PEEKNPYKE VYTDMWVEPE AAAYAPPPA
451 KKPRKSTTEK PKVKEIIDER TRERLVYEV R QKCRNIEDIC ISCGSLNVTI
501 EHPLFIGGMC QNCKNCFLEC AYQYDDDGYQ SYCTICCGGR EVLMCGNNNC
551 CRCFCVECVD LLVPGAAQA AIKEDPWNCY MCGHKGTYCL LRRREDWPSR
601 LQMFFANNHD QEFDPPKVYP PVPAEKRKPI RVLSLFDGIA TGLLVLKD LG
651 IQVDRYIASE VCEDSI TVGM VRHQGKIMYV GDVRSVTQKH IQEWGPFDLV
701 IGGSPCNLDS IVNPARKGLY EGTGRLFFEF YRLLHDARPK EGDDRPFFWL
751 FENVVAMGVS DKRDISRFLE SNPVMIDAKE VSAAHRARYF WGNLPGMNRP
801 LASTVNDKLE LQECLEHGRI AKFSKVRTIT TRSNSIKQCK DQHFPVFMNE
851 KEDILWCTEM ERVFGFPVHY TDVSNMSRLA RQRLLGRSWS VPVIRHLFAP
901 LKEYFACV*

FIG. 2A

Mouse Dnmt3b1 Protein

1 MKGDSRHLNE EEGASGYEEC IIVNGNFSQ SSDTKDAPSP PVLEAICTEP
51 VCTPETRGRR SSSRLSKREV SSLLNYTQDM TGDGDRDDEV DDGNGSDILM
101 PKLTRETKDT RTRSESPAVER TRHSNGTSSL ERQRASPRIT RGRQCRHHVQ
151 EYPVEFPATR SRRRRASSSA STPWSSPASV DFMEEVTPKS VSTPSVDLSQ
201 DGDQEGLMDTT QVDAESRDGD STEYQDDKEF GIGDLVWGKI KGF SWWPAMV
251 VSWKATSKRQ AMPGMRWVQW FGDGKFSEIS ADKLVALGLF SQHFNLATFN
301 KLVSYRKAMY HTLEKARVRA GKTFSSSPGE SLEDQLKPML EWAHCCFKPT
351 GIEGLKPNKK QPVVNNSKVR RSDSRNLEPR RRENKSRRRT TNDSAASESP
401 PPKRLKTNSY GGKDRGEDEE SRERMASEVT NNKGKLNEDRC LSCGKKNPVS
451 FHPLFEGGLC QSCRDRFLEL FYMYDEDGYQ SYCTVCCEGR ELLLCSNTSC
501 CRCFCVCECLE VLVGACTAED AKLQEPWSCY MCLPQRCHGV LRRRKDWNNMR
551 LQDFFTTDPD LEEFEPPKLY PAIPAAKRRP IRVLSLFDG I ATGYLVLKEL
601 GIKVEKYIAS EVCAESIAVG TVKHEGQIKY VNDVRKITKK NIEEWGPFDL
651 VIGGSPCN DL SNVNPKGL YEGTGRFFE FYHLLNYTRP KEGDNRPFFW
701 MFENVVAMKV NDKKDISRFL ACNPVMIDAI KVSAAHRARY FWGNLPGMNR
751 PVMASKNDKL ELQDCLEFSR TAKLKKVQT I TTKSNSIRQG KNQLFPVVMN
801 GKDDVLWCTE LERIFGFPAH YTDXSNMGRG ARQKLLGRSW SVPVIRHLFA
851 PLKDYFACE*

FIG. 2B

Human DNMT3A Protein

1 MPAMPSSGPG DTSSAAERE EDRKDOEEQE EPRGKEERQE PSTTARKVGR
51 PGRKRKHPPV ESGDTPKDPA VISKSPSMAQ DSGASELLPN GDLEKRSEPQ
101 PEECSPAGGQ KGGAPAEGEC AAETLPEASR AVENGCTPK EGRCAPAEAG
151 KEQKETNIES MKMECSRGRRL RGGLGWESSL RQRPMPRLTF QAGDPYYISK
201 RKRDEWLARW KREAEEKKAKV IAGMNAVEEN QGPGESQKVE EASPPAVQQP
251 TDPASPTVAT TPEPVGSDAG DKNATKAGDD EPEYEDGRGF GIGELWGKL
301 RGF SWWPGRIV VSWWMTGRSR AAEGTRWMW FGDCKFSVVC VEKLMPLOSSF
351 CSAFHQATYN KOPMYRKAIIY EVLQVASSRA GKLFPVCHDS DESDTAKAVE
401 VQNKPMLIEWA LGGFQPSGPK GLEPPEEKN PYKEVYTDWW VEPEAAAYAP
451 PPPAKKPRKS TAEKPKVKEI IDERTRERLV YEVHQKCRNI EDICISCGSL
501 NVTLEHPLFV GGMCQNCKNC FLECAYQYDD DGYQSYCTIC CGGREVLMCG
551 NNNCCRCFCV ECVDLLVPGC AAQAAIKEDP WNCYMCGHKG TYGLLRRRED
601 WPSRLQMFFA NNHDQEFDPP KVYPPVPAEK RKPIRVLSLF DGIATGLLVL
651 KDLGIQVDRY IASEVCEDSI TVGMVRHQGK IMYVGDVRSV TQKHIQEWP
701 FDLVIGGSPC NDLSIVNPAR KGLYEGTGRL FFEFYRLLHD ARPKEGDDRP
751 FFWLFEENVVA MGVSQDKRDIS RFLESNPVMI DAKEVSAAHR ARYFWGNLPG
801 MNRPLASTVN DKLELQECLE HGRIAKFSKV RTITTRSNSI KQGKDQHFPV
851 FMNEKEDILW CTEMERVFGF PVHYTDVSNM SRLARQRLLG RSWSVPVIRH
901 LFAPLKEYFA CV*

FIG. 2C

Human DNMT3B1 Protein

1 MKGDTRHLNG EEDAGGREDS ILVNGACSDQ SSDSPPILEA IRTPEIRGRR
51 SSSRLSKREV SSLLSYTQDL TGDGDEDGD GSDTPVMPKL FRETRTRSES
101 PAVRTRNNNS VSSRERHRPS PRSTRGRQGR NHVDESPVEF PATRSLRRRA
151 TASAGTPWPS PPSSYLTIDL TDDTEDTHGT PQSSSTPYAR LAQDSQQGGM
201 ESPQVEADSG DGDSEYQDG KEFGIGDLW GKIKGFSWWP AMVVSWKATS
251 KRQAMSGMRW VQWFGDGKFS EVSADKLVAL GLFSQHFNL A TFNKLVSYRK
301 AMYHALEKAR VRAGKTFPSS PGDSLEDQLK PMLEWAHCCF KPTGIEGLKP
351 NNTQPVNKS KVRRAGSRKL ESRKYENKTR RRTADDSATS DYCPAPKRLK
401 TNCYNNGKDR GDEDQSREQM ASDVANNKSS LEDGCLSCGR KNPVSFHPLF
451 EGGLCQTCD RFLELFYMYD DDGYQSYCTV CCEGRELLLC SNTSCCRCFC
501 VECLEVLVGT GTAAEAKLQE PWSCYMCLPQ RCHGVLRKK DWNVRLQAFF
551 TSDTGLEYE A PKLYPAIPAA RRRPIRVLSL FDGIATGYLV LKELGIVGK
601 YVASEVCEES IAVGTVKHEG NIKYVNDVRN ITKKNIEEWG PFDLVIGGSP
651 CNDLSNVNPA RKGLYEGTGR LFFEFYHLLN YSRPKEGDDR PFFWMFENVV
701 AMKVGDKRDI SRFLECNPM IDAIKVSAAH RARYFWGNLP GMNRPVIA SK
751 NDKLELQDCL EYNRIAKLKK VQTITTKSNS IKQGKNQLFP VMNGKEDVL
801 WCTELERIFG FPVHYTDVSN MGRGARQKLL GRSWSVPVIR HLFAPLKDYF
851 ACE*

FIG. 2D

Dnmt3a	1	MPSSGPGDTSSSLEREDDRKEGEEQEENRGKEERQEPSATARKVGRPGR	50
Dnmt3a	51	KRKHPPVESSDTPKDPAVTTKSQPMAQDSGPSD....LLPNGDLEKRSEP	96
		. . . : : : : . .	
Dnmt3b	1MKGDSRHLNEEGASGYEECIIVNGNFSQSDQSSD	33
Dnmt3a	97	QPEEGSP....AAGQKGGAPAEAGEGETTPPEAS.RAVENGCCVTKE..GR	139
		: 	
Dnmt3b	34	TKDAPSPPVLEAICTEPVCTPETRGRRSSSRLSKREVSSLNYTQDMTGD	83
Dnmt3a	140	G.....ASAGEG.....KEQKQTNIIESMKMEGSRGRRLRGGLGWESSLRQ	178
		. : : :	
Dnmt3b	84	GDRDDEVDDGNNGSDILMPKLTRTETKDTTRTRSESPA VRTRHSNGTSSLERQ	133
Dnmt3a	179	RPMPLRTFQAGDPYYISKRKRDEWLARWKREAEEKAKVIAVMNAVEENQA	228
		: : : : : : : : :	
Dnmt3b	134	RASPRITRGRQGRHHV....QEYPVEFPATRSRRRASSSASTPWSSPA	178
Dnmt3a	229	SGESQKVVEEASPPAVQQPTDPASPTVATTPEPVGGDAGDKNATKAADDEP	278
		: . .	
Dnmt3b	179	SVDF..MEEVTPKSVSTP....SVDLSDQGDQEGMDTTQVDAESRDG DST	222
Dnmt3a	279	EYEDGRGFGIGELVWGKLRGFSWWPGRI VSWWMTGRSRAAEGTRWVMWFG	328
		: : :	
Dnmt3b	223	EYQDDKEFGIGDLVWGKIKGFSWWPAMVVSWKATSKRQAMPGMRWVQWFG	272
Dnmt3a	329	DGKFSVCVCKLMPPLSSFCASFHQATYNKQPMYRKAIYEVLQVASSRAGK	378
		: :	
Dnmt3b	273	DGKFSEISADKLVALGLFSQHFNLATFNKLVSYRKAMYHTLEKARVRA GK	322
Dnmt3a	379	LFPACHDSDESDSGKAVEVQNQKQMI EWALGGFQPSGPKGLEPPEEK..N	426
		.	
Dnmt3b	323	TF.....SSSPGESLEDQLKPMLEWAHGGFKPTGIEGLKPNKKQPVN	365
Dnmt3a	427	PYKEVYTDMW.VEP.....EAAAYAPPPPAAKKPRKSTTEKPK	462
		. . :	
Dnmt3b	366	KS KVRRSDSRNLEPRRRRENKSRRRTNDSAASESPPP KRLKTNSYGGKDR	415

FIG.3A-1

Dnmt3a	463	VKEIIDERTRERLVYEVROKCRNIEDICISCGSLNVTLHPFFIGGMQCN	512
	.	. : - : : - : .	
Dnmt3b	416	GE...DEESRERMASEVTNNKGNLEDRCLSCGKKNPVSFHPLFEGGLCQS	462
Dnmt3a	513	CKNCFLECAYQYDDDGYQSCTICCGGREVLMCGNNNCCRCFCVECDLL	562
	. : : . : . ..		
Dnmt3b	463	CRDRFLELFYMYDEDGYQSCTVCCEGRELLCSNTCCRCFCVECLEVL	512
Dnmt3a	563	VGPGAAQAAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHD.Q	611
	: : . . : . . . :		
Dnmt3b	513	VGAGTAEDAKLQEPWSCYMCQPQRCHGVLRRRKDWNMRLQDFFTDPDLE	562
Dnmt3a	612	EFDPPKVYPPVPAEKRKPIRVLSLFDFIATGLLVLKDLGIQVDRYIASEV	661
	: .. : : : . :		
Dnmt3b	563	EFEPPKLYPAIPAACKRPIRVLSLFDFIATGYLVLKELGIKVEKYIASEV	612
Dnmt3a	662	CEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWPFDLVIGGSPCNLSI	711
	: : . : . :		
Dnmt3b	613	CAESIAVGTVKHEGQIKYVNDVRKITKKNIEEWGPFDLVIGGSPCNLSN	662
Dnmt3a	712	VNPARKGLYEGTGRFFFYRLLHDARPKEGDDRPFFWFENVVAMGVSD	761
	. . . : .		
Dnmt3b	663	VNPARKGLYEGTGRFFFYHLLNYTRPKEDGNRPFFWMFENVVAMKVND	712
Dnmt3a	762	KRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLEL	811
	: . ..		
Dnmt3b	713	KKDISRFLACNPVMIDAIVSAAHRARYFWGNLPGMNRPMASKNDKLEL	762
Dnmt3a	812	QECLEHGRIAKFSKVRTITRSNSIKQGKDQHFPVFMNEKEDILWCTEME	861
	: . : : . : :		
Dnmt3b	763	QDCLEFSRTAKLKKVQTITKSNSIRQGKNQLFPVVMNGKDDVLWCTEL	812
Dnmt3a	862	RVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV*	909
	: : :		
Dnmt3b	813	RIFGFPAHYTDVSNMGRGARQKLLGRSWSVPVIRHLFAPLKDYFACE*	860

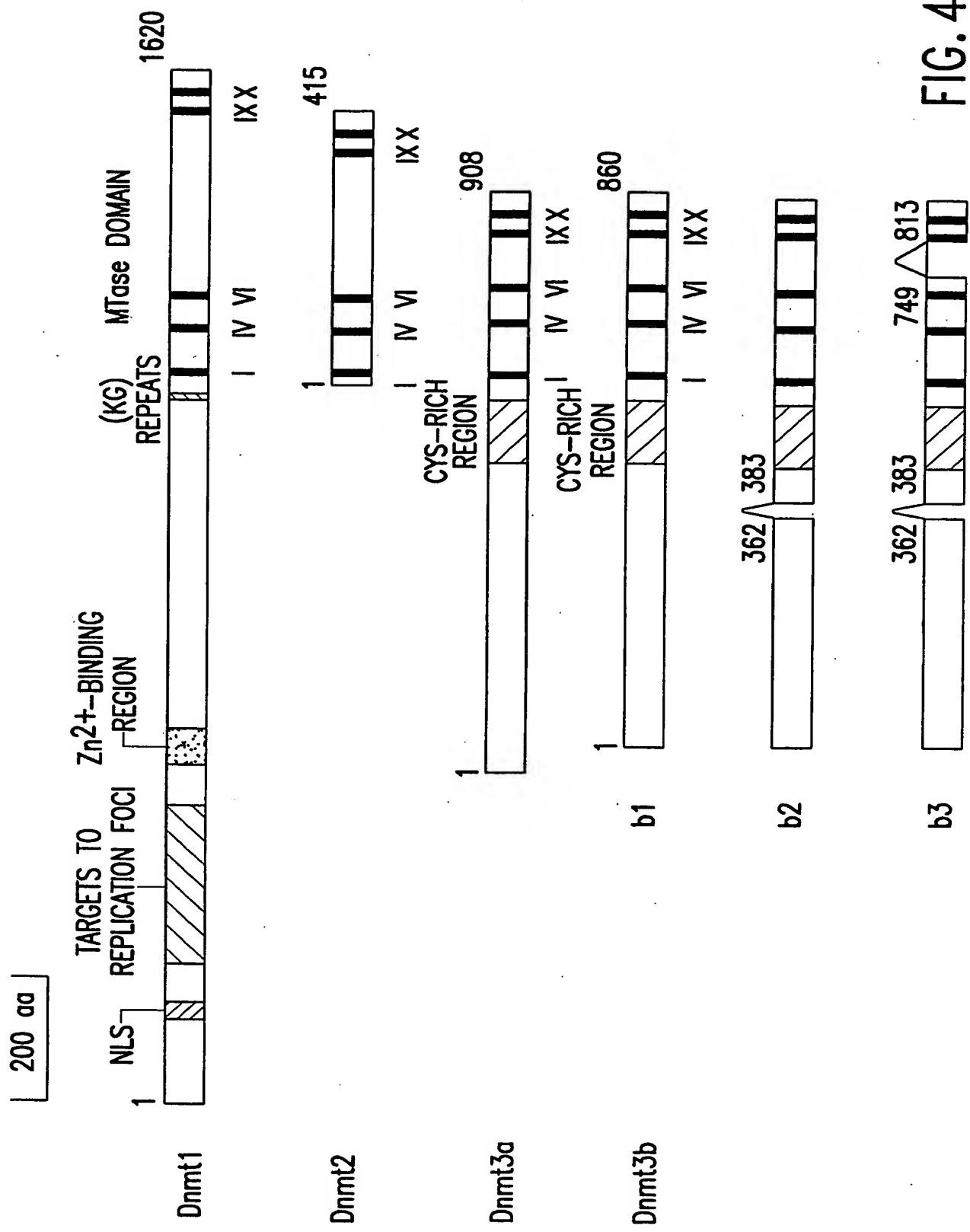
FIG.3A-2

DNMT3A	1	MPAMPSSPGPDTSSAAEREEDRKDGEEQEEPGRGKEERQEPSTTARKVGR
DNMT3A	51	PGRKRKHPPVESGDPKDPAVISKSPSMAQDSGASELLPNGDLEKRSEPQ
DNMT3B	1MKGDRHLNGEEDAGGREDSILVNGACSDQSSDSP
DNMT3A	101	PEEGSPAGGQKGGAPAEGEGAAETLPEASRAVENGCCTPKERGRGAPAEAG
DNMT3B	36	PILEAIRTPEIFRGWASSRLSKREVSSLSSYTQDLTGDGDEDGDGS DTP
DNMT3A	151	KEQKETNIESMKMEGSRGRRLRGGLGWESSLRQRPMPLRTFQAGDPYYISK
DNMT3B	86	VMPKLFRETTRTRSESPAVRTRNNNSVSSRERHRPSPRSTRGRQGRNHVDE
DNMT3A	201	RKRDEWLARWKREAEEKKAKVIAGMNAVEENQGPGESQKVEEASPPAVQQP
DNMT3B	136	SPVEFPATRSLRRRATASAGTPWPSPPSSYLTIDLTDDEDTH..GTPQS
DNMT3A	251	TDPASPTVATTPEPVGSAGDKNATKAGDDEPEYEDGRGFGIGELVGKL
DNMT3B	184	SSTPYARLAQDSQQGGMESPQVEADSGDGSSEYQDGKEFGIGDLVGKI
DNMT3A	301	RGFSWWPGRIVSWWMTGRSRAAEGRTRWMWFBDGKFSVVCKMLPLSSF
DNMT3B	234	KGFSWWPAMVWSWKATSKRQAMSGMRWVQWFGDGKFSEVSADKLVALGLF
DNMT3A	351	CSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPVCHDSDESDTAKAVE
DNMT3B	284	SQHFNLATFNKLVSYRKAMYHALEKARVRAGKTFP.....SSPGDSLE
DNMT3A	401	VQNKPMLIEWALGGFQPSGPKGLEP....PEEEKNPYKEVYTDMWVE....
DNMT3B	327	DQLKPMLEWAHGGFKPTGIEGLKPNNTQPVVNKSVRAGSRKLESRKYE
DNMT3A	443PEAAAYAPPPPAKKPRKSTAEPKVKEIIDERTRERLVYEVRO
DNMT3B	377	NKTRRRRTADDTSATSDYCPAPKRLKTNCYNNKGDRGDEDQSREQMASDVAN

FIG. 3B-1

DNMT3A	486	KCRNIEDICISCGSLNVTLEHPLFVGGMQNCNCFL ECAYQYDDDGYQS
		.:: : : .:
DNMT3B	427	NKSSLEDGCLSCGRKNPVSFHPLFEGGLCQTCRDRFLELFYMYDDDGYQS
		: . : . . : .
DNMT3A	536	YCTICCGGREVLMCGNNNCRCFCVECVDLLVGPGAAQAAIKEDPWNCYM
		: . : . . : .
DNMT3B	477	YCTVCCEGRELLLCNTSCCRCFCVECLEVLVGTGTAEEAKLQEPWSCYM
		: . : . . : .
DNMT3A	586	CGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPKVYPPVPAEKRKPIR
		. : . .. : . : :
DNMT3B	527	CLPQRCHGVLRKKDWNVRLQAFFTSDTGLEYEAPKLYPAPIAARRRPIR
		: . : . . : .
DNMT3A	636	VLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVG
		: . : . . : .
DNMT3B	577	VLSLFDGIATGYVLKELGIKVGVKYVASEVCEESIAVGTVKHEGNIKYVN
		: . : . . : .
DNMT3A	686	DVRSVTQKHIQEWPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFEEFY
		: . : . . : .
DNMT3B	627	DVRNITKKNIEEWGPFDLVIGGSPCNDLSNVNPARKGLYEGTGRLFEEFY
		: . : . . : .
DNMT3A	736	RLLHDARPKEGDDRPFFWL FENVVAMGVSDKRDISRFLESNPVMIDAKEV
		. . . : . .
DNMT3B	677	HILLNYSRPKEGDDRPFFWMFENVVAMKVGDKRDISRFLECNPVMIDAIKV
		: . : . . : .
DNMT3A	786	SAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITT
		: . : . . : .
DNMT3B	727	SAAHRARYFWGNLPGMNRPVIA SKNDKLELQDCLEYNRIAKLKKVQTITT
		: . : . . : .
DNMT3A	836	RSNSIKQGKDQHFPFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLAR
		: . :
DNMT3B	777	KSNSIKQGKNQLFPVVMNGKEDVLWCTELERIFGFPVHYTDVSNMGRGAR
		: . : . . : .
DNMT3A	886	QRLLGRSWVPVIRHLFAPLKEYFACV*
		: :
DNMT3B	827	QKLLGRSWVPVIRHLFAPLKDYFACE*

FIG.3B-2



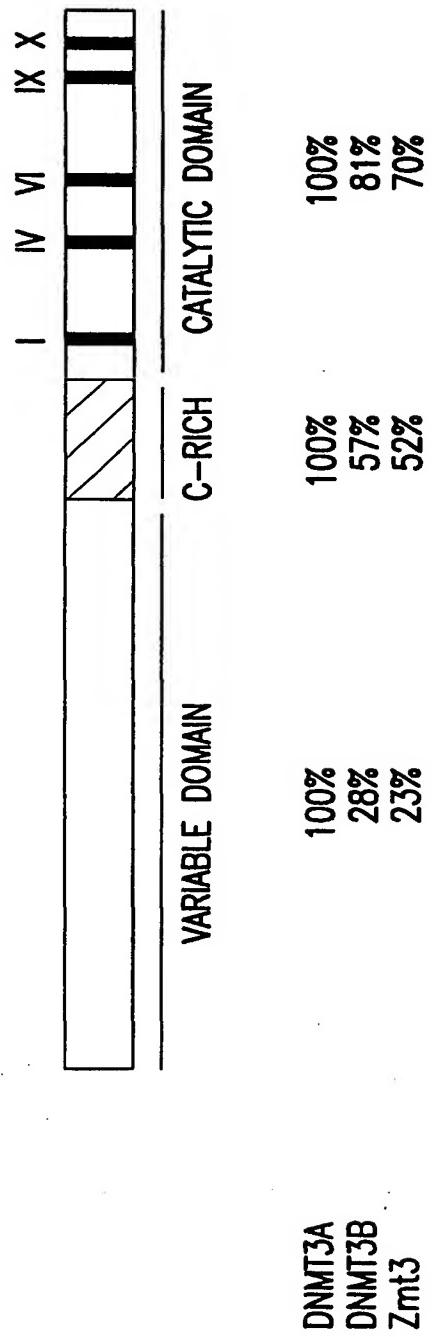


FIG. 4B



FIG. 4C

FIG. 4D

Exon1 (>=90bp) CGGCAGgtgaggccccgggg.intron(17618bp).tggcttcctccacaggaaAGC
Exon2 (148bp) TCAGAGgtggctggcagtgg.intron(887bp) .CTGTTTCCTCTACAGGCCGAA
Exon3 (62bp) ACACAGgttatggttctgtctc.intron(3343bp).tggttccctataaaaggACTTGG
Exon4 (102bp) CCAGCTgttaaagttagccacacc.intron(1642bp).ctctcttgcttcttagtCCGA
Exon5 (125bp) ACCAGGtttgttcccaggatg.intron(602bp).tccctctgtccacaggTCCCTG
Exon6 (222bp) TATCAGgttatggccgaggg.intron(1403bp).tgggttttccaggATGGG
Exon7 (159bp) TCCGAGgttgagtccggggaaag.intron(2588bp).gtcttttcttttagGTCTCT
Exon8 (108bp) CTGGAGgttaaacatggatgag.intron(917bp).actctggctttgcagaAAAGCT
Exon9 (145bp) AACCAGgtggaaatggatgtccc.intron(765bp).ttttccctcaaaaggTTGTTA
Exon10 (60bp) AATAACGgttattttcttcgtt.intron(1813bp).attacctttcacaggAAACA
Exon11 (126bp) GCGGAGgttgattttgggtac.intron(115bp).tctttttctcaatagaACAAA
Exon12 (45bp) TGGAAGgttaaacgttctcc.intron(1095bp).ctgtttttcttacagATGGCT
Exon13 (80bp) TGCCGGgttaagtcttcact.intron(417bp).ctctctggctgtccaggATCGC
Exon14 (113bp) CTGCCGGgttagcactggggcc.intron(1160bp).tgccactgggtccaggGTGTT
Exon15 (184bp) GAATAACgttaaggccacaggctc.intron(600bp).ttccttacctggcaggAAAGCC
Exon16 (85bp) CGACAGgtgaggttcggggaac.intron(824bp).ctctggccccacaggCTAC
Exon17 (146bp) AAAAATgtgaggcagtctgt.intron(536bp).gtctcttctttcagATTGAA
Exon18 (91bp) TGTATGgttagcatcccttctc.intron(352bp).cttttcttagcacaaggGGTA
Exon19 (149bp) CTGGAGgtgaggaaatctggg.intron(958bp).tctttctcccccacaggTGTAAAT
Exon20 (86bp) GAACAGgttaacaagggtct.intron(2867bp).tttggctgttccaggGCCGT
Exon21 (70bp) GCCAAGgttaaaaaggatcacag.intron(801bp).cattttgttctccaggTTAAAG
Exon22 (119bp) CGAAAGgttagcaaggctgca.intron(1434bp).ctccggtaaccccaaggGATCT
Exon23 (1585bp)

I DNMT1 Dnmt1 MET1 (Ath) Masc1 Masc2 Dnmt2 M.Spr DNMT3A Dnmt3a DNMT3B Dnmt3b Zmt3 consensus	IV DVS GCGGL SEG FHQAG DV S GG G GL SE GFHQAG DIFAG GGGL SHGLKKAG DT FGGGV SLGARQAG DIFAG GGGL TGLDLSG ELYSGIGGMHHALRESH SLFSGIGAFEAA LRNIG SLFDGIAT GLL V KDLG SLFDGIAT GLL V KDLG SLFDGIAT GLL V KELG SLFDGIAT GLL V KELG SLFDGIAT GLV L RDLG -F-G - consensus	V DWEMLC GGPP CQGFSGMNR DWEMLC GGPP CQGFSGMNR QWDFTING GGPP CQGFSGMNR HDIILHLS P CQTFSRAHT EVDFYGG PP CQGFSGMNR SFNM11MS P CQP FTRIGL EFDLV GGSP CQSFSVAGH PFDLV GGSP CNDLSIVNP PFDLV GGSP CNDLSIVNP PFDLV GGSP CNDLSIVNP PFDLV GGSP CNDLSIVNP PFDLLIGGS PC N PC -GG- - consensus	VI YRP RF FFLLE ENVNFVFSFKR YRP RF FFLLE ENVNFVFSYRR FRPRYF FLLE W RTFVFSFNK VRPRLFTVEET FGIMDRQ YKPR RF V U LENW M G LI TKL KLPKY T ILEW M G FEV SST KOPKF V FEW M G LI NHDK DRPF FWL FEN W AMGVSDK DRPF FWL FEN W AMGVSDK DRPF FWL FEN W AMGVSDK DRPF FWL FEN W AMGVSDK DRPF FWL FEN W AMGVSDK -P-F- - consensus
---	---	---	---

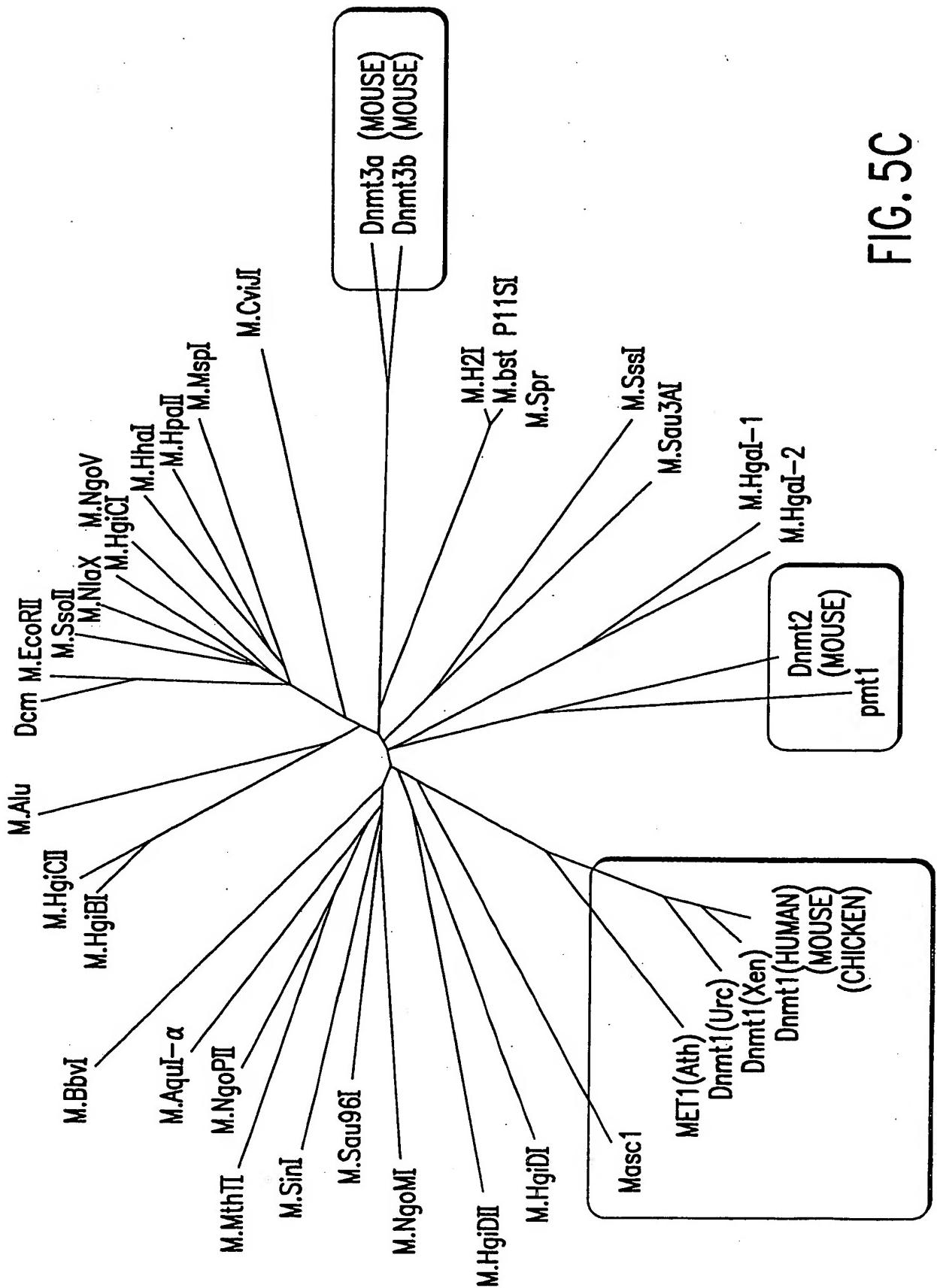
IX RVSVRE CA RSQ GFP RVSVRE CA RSQ GFP RILT V RE CA RSQ GFP RKFTV RELACI Q GFP RVYT V RELARAQ GFP RYFTPKETIANLQ GFP RRLTPLECFRLQAFD DILWC TEM ERVF GFP DILWC TEM ERVF GFP DVLWC CTE LERIF GFP DVLWC CTE LERIF GFP DHIWITTELEK IF G GFP -E-R-GFP - consensus	X LGNILD K H R QVGNAV PP PLAKAIG FFGNILD R H R QVGNAV PP PLAKAIG FAGNINHKH R QIGNAV PP PLAFALG FVGTLTDK R RIIGNAV PP PLSAAIM GLGGVKKMHRNIGNAV PP LGEQIG EKT TV KQR Y RLIGNSLN V HWAKLL AGISNSQLYKQ T GNISIT T VL E SIF SNMSRLARQ R LLGRSWS S VP V IRHLF SNMSRLARQ R LLGRSWS S VP V IRHLF SNMGRGAROK K LLGRSWS S VP V IRHLF KSMGRPQRQRV L GRSWS S VP V IRHLF -R-G- - consensus
---	---

FIG. 5A

DNMT3A	EDICISCG.....	SLNVTLEHPLFVGGMQCNCKNCFL	EDICISCG.....	EDICISCG.....
Dmrt3a	EDICISCG.....	FECAYQYDDDGYQS	SLNVTLEHPLFVGGMQCNCKNCFL	EDICISCG.....
DNMT3B	EDGCLSCG.....	EDGCLSCG.....	EDGCLSCG.....	RKNPVSFHPLFEGGLCQT
Dmrt3b	EDGCLSCG.....	EDGCLSCG.....	EDGCLSCG.....	TCRDRFLEFYMYDDG
Zmt.3	EDRCLSCG.....	EDRCLSCG.....	EDRCLSCG.....	EDGQSYCT
ATRX Human	EDFCLSCG.....	EDFCLSCG.....	EDFCLSCG.....	KKNPVSFHPLFEGGLCQS
ATRX Mouse	IVSCTACGQQVNHFQKDSIYRPSLQV	IVSCTACGQQVNHFQKDSIYRPSLQV	IVSCTACGQQVNHFQKDSIYRPSLQV	CRFLEFLFYMDE
Consensus	C C	C C	C C	EDGYQS
DNMT3A	IICCGGREVLMCGNNNCCRCFC	IICCGGREVLMCGNNNCCRCFC	IICCGGREVLMCGNNNCCRCFC	IICCGGREVLMCGNNNCCRCFC
Dmrt3a	VECDLLVGPGAAQAAIKE.	VECDLLVGPGAAQAAIKE.	VECDLLVGPGAAQAAIKE.	VECDLLVGPGAAQAAIKE.
DNMT3B	DPMWCYMCGHKG	DPMWCYMCGHKG	DPMWCYMCGHKG	DPMWCYMCGHKG
Dmrt3b	VCCEGRELLCSNTSCRCFC	VCCEGRELLCSNTSCRCFC	VCCEGRELLCSNTSCRCFC	VCCEGRELLCSNTSCRCFC
Zmt.3	VECLEVLVGAGTAEDAKLQ.	VECLEVLVGAGTAEDAKLQ.	VECLEVLVGAGTAEDAKLQ.	VECLEVLVGAGTAEDAKLQ.
ATRX Human	EPWSCYMCILPQRC	EPWSCYMCILPQRC	EPWSCYMCILPQRC	EPWSCYMCILPQRC
ATRX Mouse	WCAEGGNLICC..DFCHNAFKKKCILRNLRKELSTIMDENNO	WCAEGGNLICC..DFCHNAFKKKCILRNLRKELSTIMDENNO	WCAEGGNLICC..DFCHNAFKKKCILRNLRKELSTIMDENNO	YCICHPEPL
Consensus	C C	C C	C C	YCICHPEPL

FIG. 5B

FIG. 5C



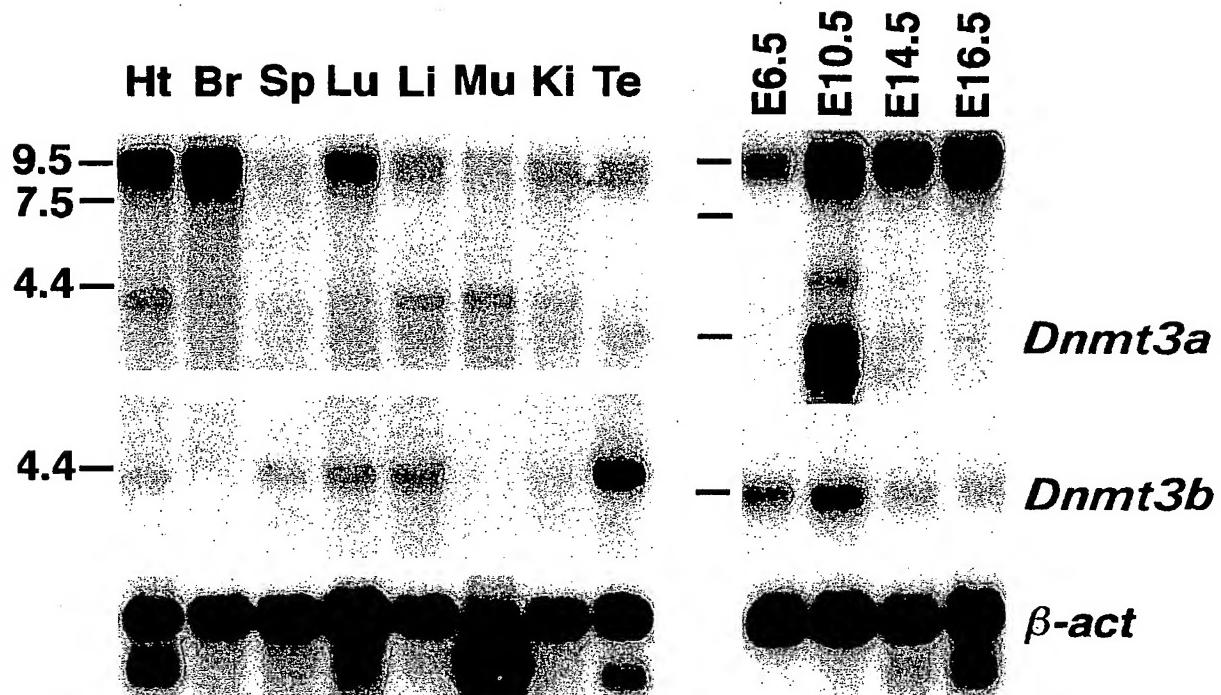


FIG.6A

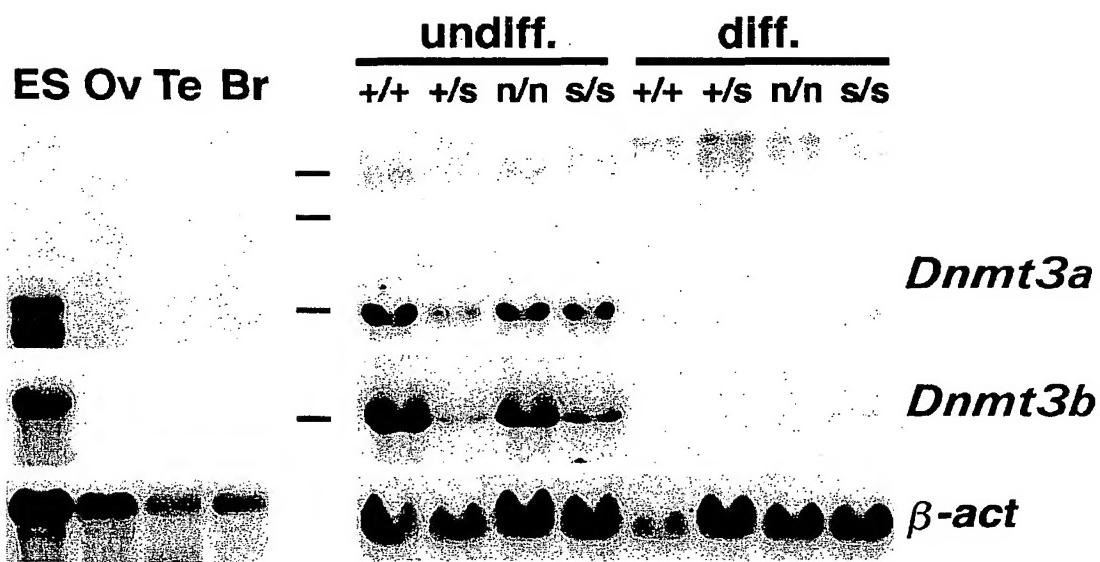
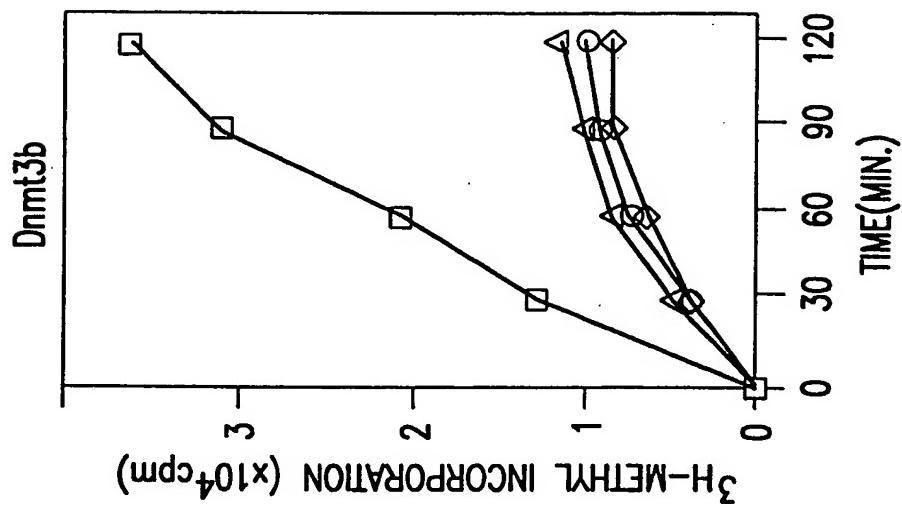


FIG.6B



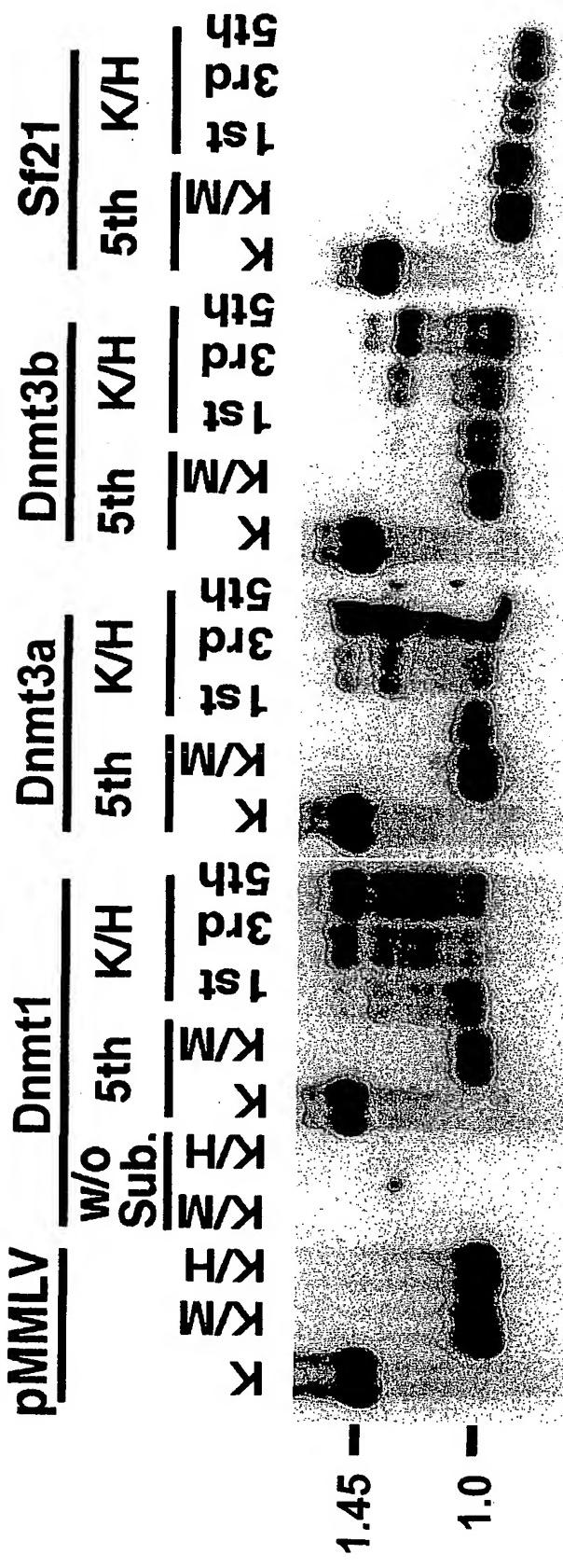


FIG. 7D

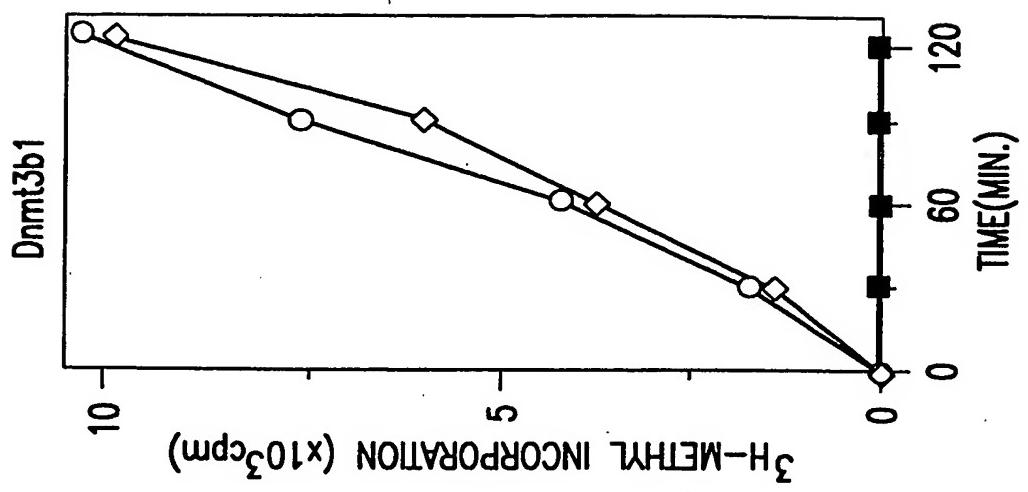


FIG. 8C

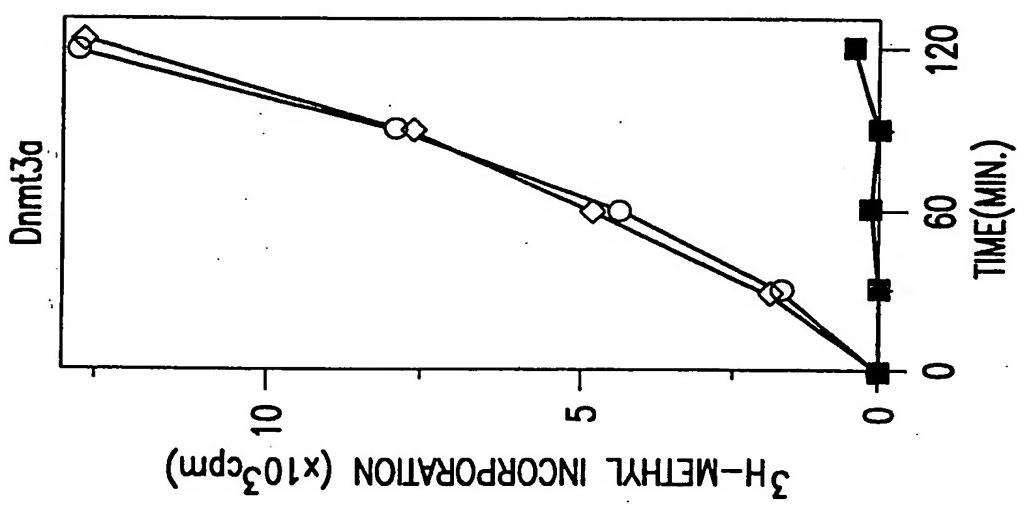


FIG. 8B

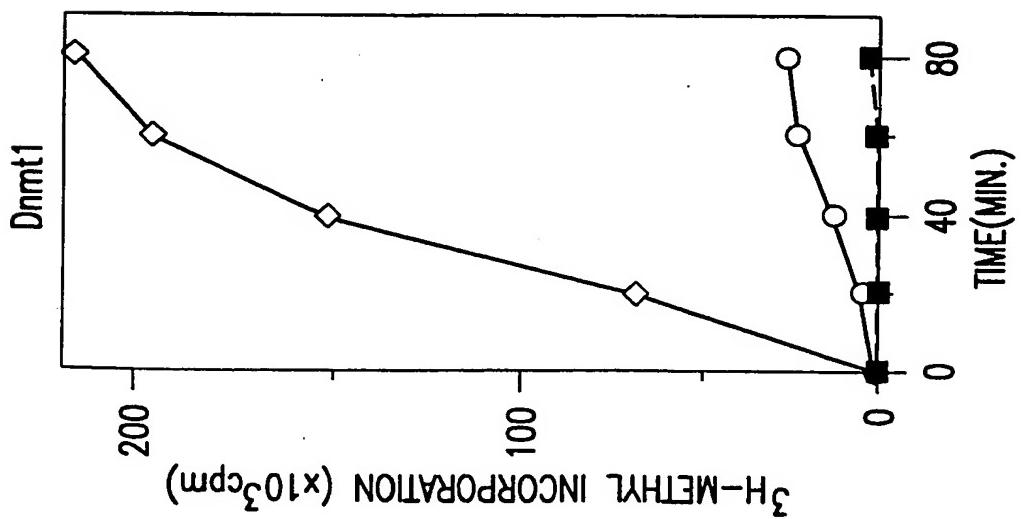


FIG. 8A

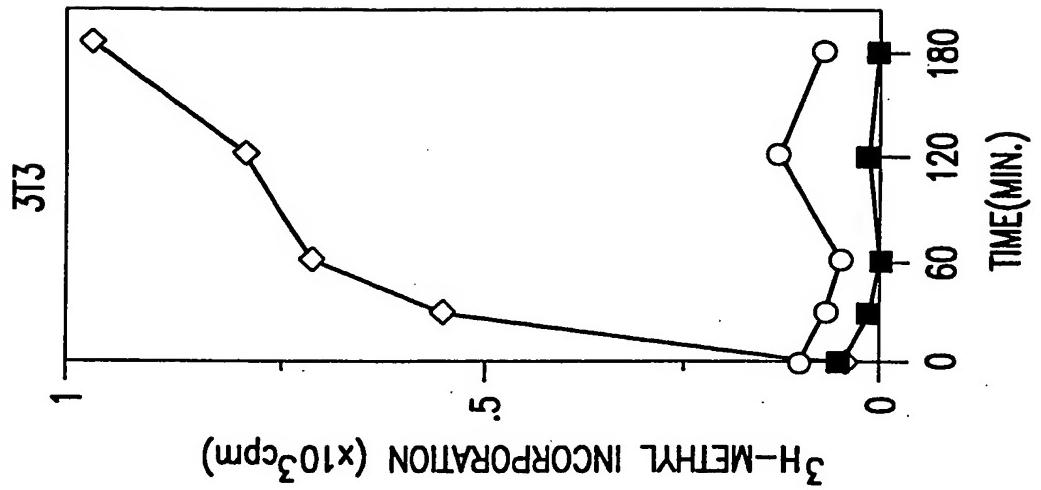


FIG. 8E

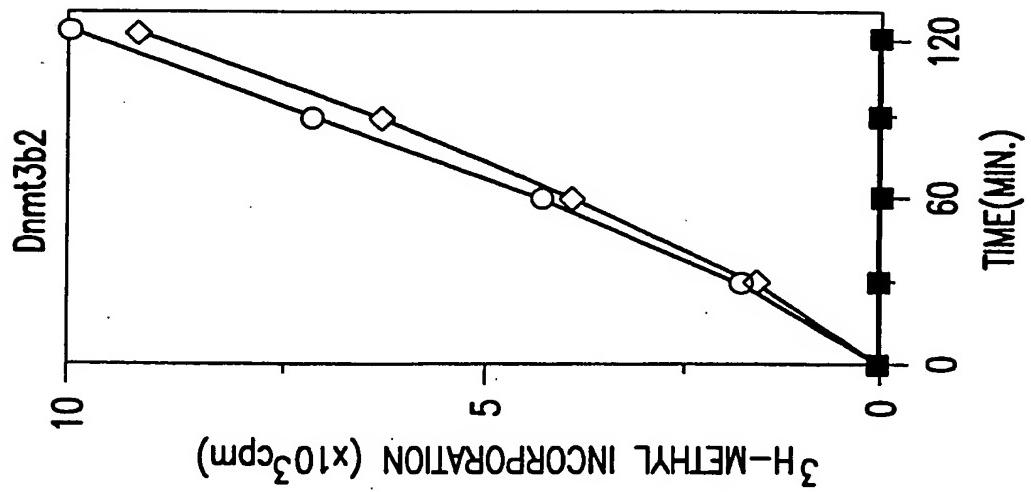
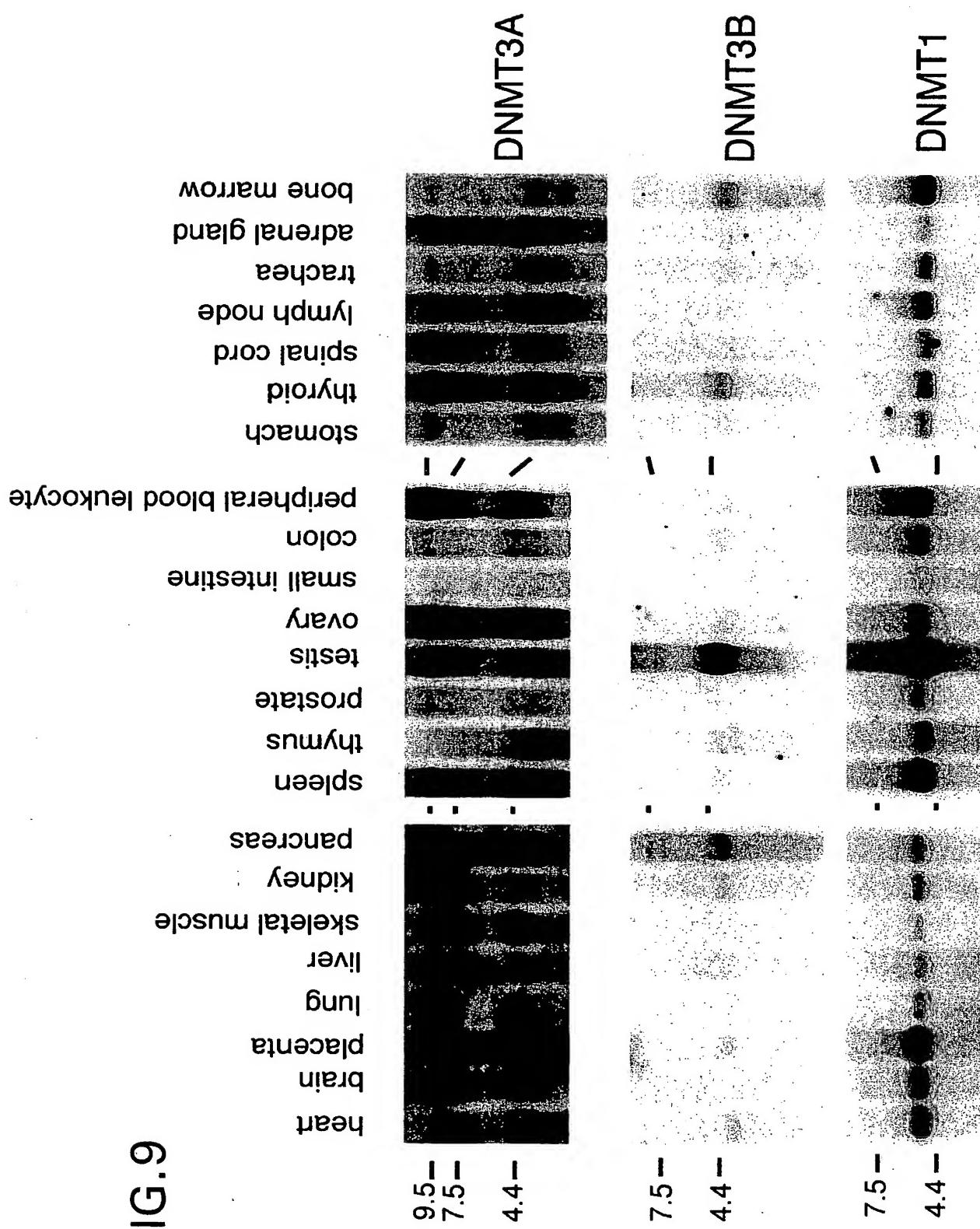


FIG. 8D

FIG. 9



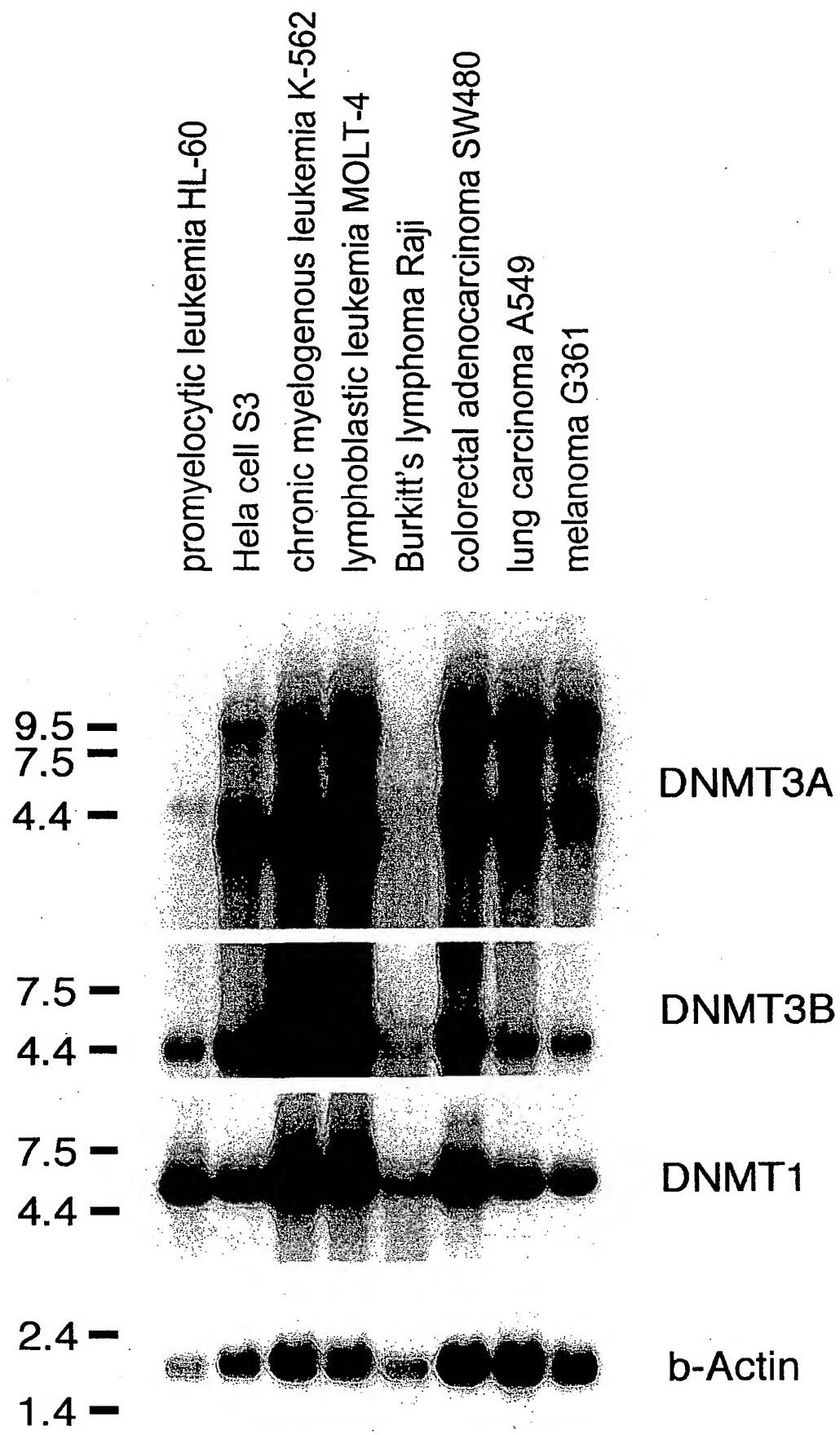


FIG. 10

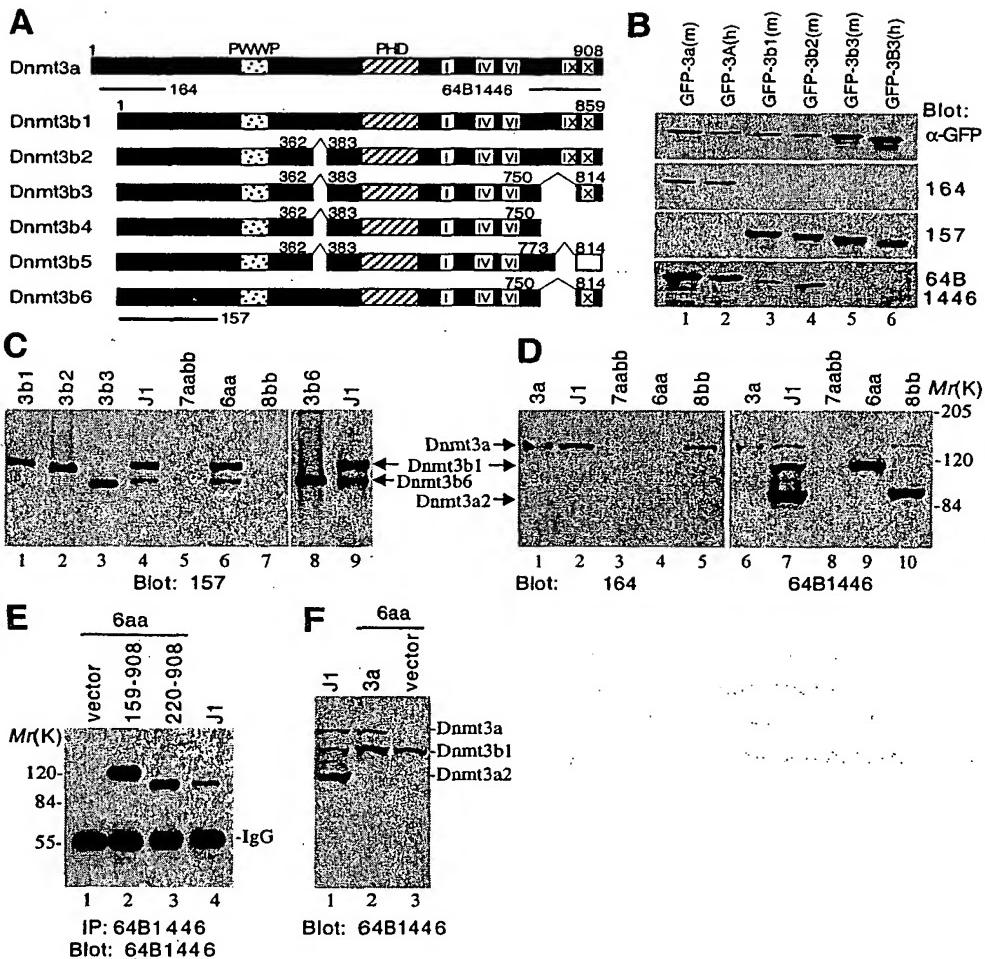


FIG. 11

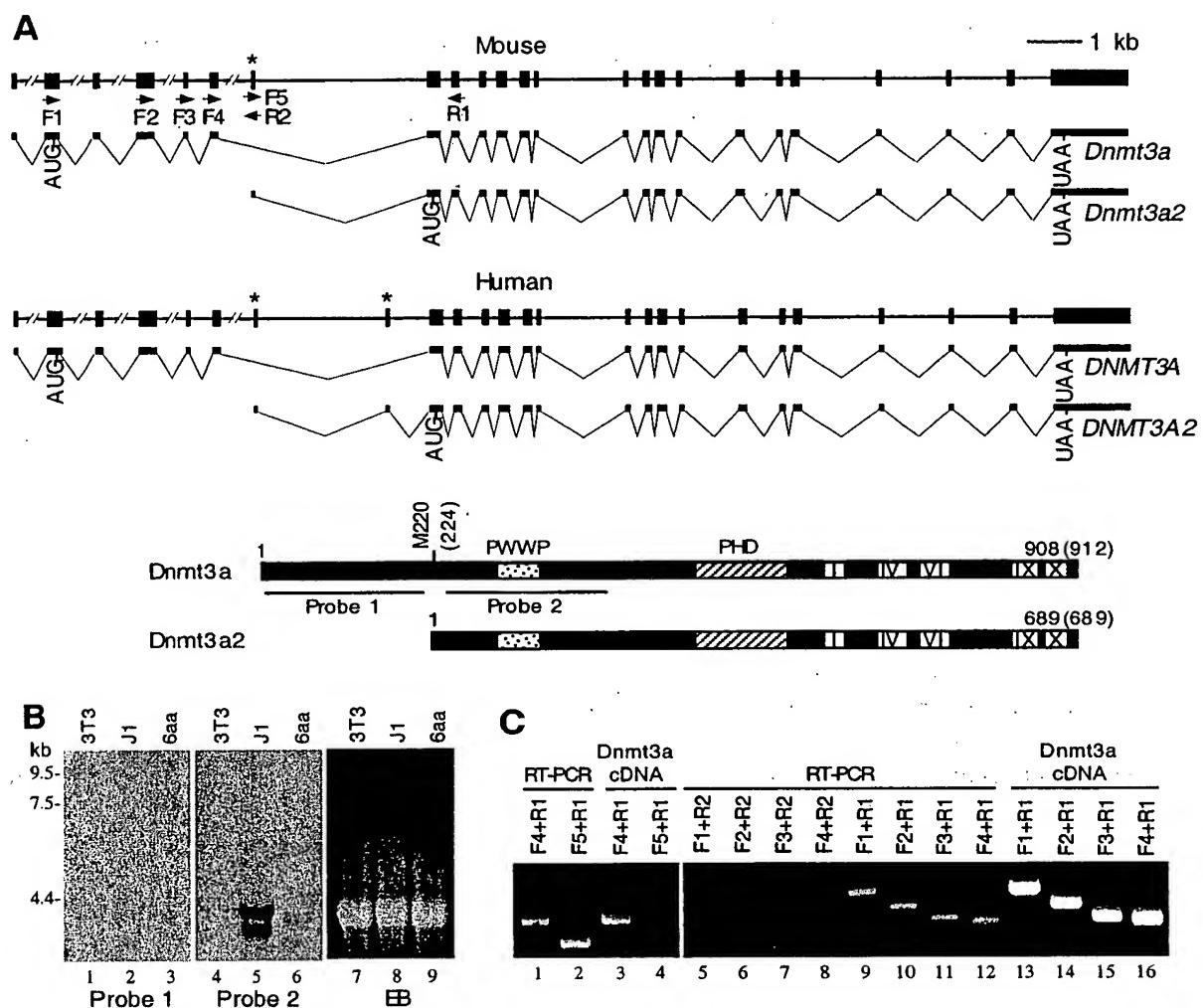


FIG. 12

Mouse Dnmt3a2 cDNA sequence:

FIG. 13A

Mouse Dnmt3a2 amino acid sequence:

1 MNAVEENQASGESQKVEEASPPAVQQPTDPASPTVATTPEPVGGDAGDKNATKAADDEP
60 EYEDGRGFGIGELVGKLRGFSWWPGRIVSWWMTGRSRAEGTRWMWFGDGKFSVVCV
119 EKLMPPLSSFCASFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPACHDSDESDSGKAVE
178 VQNQKMIEWALGGFQPSGPKGLEPPEEEKNPYKEVYTDMWVEPEAAAYAPPPPAAKKPRK
237 STTEKPKVKEIIDERTRERLVYEVROKCRNIEDICISCGLSLNVTLEHPLFIGGMQCQNC
296 NCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRCFCVECVDLLVPGAAQAAIK
355 EDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVYPPVPAEKRKPIRV
414 LSLFDGIATGLLVLKD LGIQVDRYIASEV CEDSITVGMVRHQGKIMYVGDVRSVTQKHI
473 QEWGPFDLVI GGSPCN DLSIVNPARK GLYEGT GRLFFFYRLLHDARPKEGDDR PFFWL
532 FENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVNDKL
591 ELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHF PVFMNEKEDILWCTEMERVFGFPV
650 HYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV

FIG. 13B

Human DNMT3A2 cDNA sequence:

1 ccgcggccagccccatcgcccccttcccctccccaagacgggcagctactccagagc
60 ttcaggggccggctcacacctgagcgcgactgcagagggctgacacctggcattatgg
119 ggatcctggagcgggttgtgagaaggaatgggcgcgtggatcgttagcctgaaagacgag
178 tgtgatacggctgagaagaaagccaaggcattgcaggaatgaatgtgtgaaagaaaa
237 ccaggggccccggggagtcctcagaagggtggaggaggccagccctctgtgcagcagc
296 ccactgaccccgcatccccactgtggctaccacgcctgagcccgtgggtccgatgct
355 gggacaagaatgccaccaaagcaggcgatgacgagccagagtacgaggacggccggg
414 ctttggcattggggagctggtgtggggaaactgcggggcttcctggcggccaggcc
473 gcattgtgtcttggatgacggccggagccgagcagctgaaaggcaccgcgtggc
532 atgtggttcggagacggcaaattctcagtgggtgttggagaagctgatgccgtgag
591 ctgcgtttgcagtgcgttccaccaggccacgtacaacaagcagccatgtaccgcaaag
650 cccatctacgaggctctgcagtgccagcagccgcgggaagctgttcccggtgtgc
709 cacgacagcgatgagagtgacactgccaaggccgtggaggtgcagaacaagccatgat
768 tgaatgggcctgggggcttccagcctctggccctaagggcctggagccaccagaag
827 aagagaagaatccctacaaagaagtgtacacggacatgtgggtggAACCTGAGGCAGC
886 gcctacgcaccacccaccagccaaaaagcccgaaagagcacagcggagaagcccaa
945 ggtcaaggagattattgtgagcgcacaagagagcggctgggtgtacgaggtgcggcaga
1004 agtgcggaaacattgaggacatctgcattctgtggagcctcaatgttaccctggaa
1063 caccctcttcgtggaggaatgtgcacaaaactgcaagaactgcattctggagtgtgc
1122 gtaccagtacgacgcacgcggctaccagtctactgcaccatctgtgtggggccgtg
1181 aggtgctcatgtgcggaaacaacaactgctgcaggtgtttgcgtggagtggtggac
1240 ctctggggccggggctgcccaggcattaaagaagacccctggaaactgcta
1299 catgtgcggcacaagggtacctaaccggctgtgcggcggcagaggactggccctccc
1358 ggctccagatgttctcgctaataaccacgaccaggaaattgaccctccaaaggttac
1417 ccacctgtccagctgagaagaggaagccatccgggtgtctctttgtatggaaat
1476 cgctacagggtccctgggtgtgaaaggacttggcattcaggtggaccgcataattgcct
1535 cggaggtgtgaggactccatcacgggtggcatggcggcaccaggaaagatcatg
1594 tacgtcgggacgtccgcagcgtcacacagaacatccaggagtggggccattcga
1653 tctgggtattggggcagtcctgcataatgacctctccatcgtcaaccctgctgcacagg
1712 gcctctacgagggcactggccggctttttgagttctaccgcctcctgcattgtgc
1771 cggcccaaggaggagatgatgcggccattctggctcttgagaatgtgggtggccat
1830 gggcgtagtgacaagaggacatctcgcatggatcttcgagttccaaaccctgtgatgattg
1889 atgccaagaagaagtgtcagctgcacacaggcccgtacttctgggttaacctccgg
1948 atgaacaggccgttggcatccactgtgaatgataagctggagactgcaggagtgtctgga
2007 gcatggcaggatagccaaagttcagccaaagtgaggaccattactacgaggtaaaactcca
2066 taaagcaggccaaagaccacgcatttcgtcttcatgaatgagaaagaggacatctta
2125 tgggtgcactgaaatggaaaggatattggttccactatactgacgtctccaa
2184 catgagccgttggcgaggcagagactgctggccggcatggagcgtgccagtcattcc
2243 gccacctctcgctccgctgaaggagtatttgcgtgtgtaaaggacatggggccaa
2302 actgaggttagcgacacaaagttaaacaaacaaaacacaaaacataataaaaaca
2361 ccaagaacatg

FIG. 13C

Human DNMT3A2 amino acid sequence:

1 MNAVEENQGPGESQKVEEASPPAVQQPTDPASPTVATTPEPVGSAGDKNATKAGDDEP
60 EYEDGRGFGIGELVGKLRGFSWWPGRIVSWWMTGRSRAEGTRWMWFGDGKFSVCV
119 EKLMPPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPVCHDSDESDTAKAVE
178 VQNKPMLIEWALGGFQPSGPKGLEPPPEEKNPYKEVYTDMWVEPEAAAYAPPPPAAKKPRK
237 STAEPKPKVKEIIDERTRERLVYEVROQKCRNIEDICISCGLNVTLHPLFVGGMCQNCK
296 NCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCRCFCVECVDLLVGPAAQAAIK
355 EDPWNCYMCGHKGTGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVYPPVPAEKRKPIRV
414 LSLFDGIATGLLVLKD LGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHI
473 QEWGPFDLVI GGSPCN DLSIVNPARKGLYEGTGR LF FEFYRLLHDARPKEGDDRPF FWL
532 FENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAH RARYFWGNLPGMNRPLASTVNDKL
591 ELQECLEHGRIA KFSKVRTIT TRSNSIKQGKDQHF PVFMNEKEDI LWCTEMERVFGFPV
650 HYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV

FIG. 13D

	10	20	30	40	50		
Dnmt3a2	1	ccgcccccaa	ccccaacgcc	ccctgcccct	ccccccagac	gggcagctat	50
DNMT3A2	1	ccgcccccaag	ccccatcgcc	cccttcccct	cccccaagac	gggcagctac	50
	60	70	80	90	100		
Dnmt3a2	51	ttacagagct	tc-gggccgg	ggctcacacc	tgagctgtac	tgcagagggg	100
DNMT3A2	51	ttccagagct	tcagggccgc	ggctcacacc	tgagcgcgac	tgcagagggg	100
	110	120	130	140	150		
Dnmt3a2	101	ctgcacacctg	ccttatgg--	-----	-----	-----	150
DNMT3A2	101	ctgcacacctg	ccttatgggg	atcctggagc	gggttgtgag	aaggaatggg	150
	160	170	180	190	200		
Dnmt3a2	151	-----	-----	-----	gctg	agaagaaagc	200
DNMT3A2	151	cgcgtggatc	gtagcctgaa	agacgagtgt	gatacggctg	agaagaaagc	200
	210	220	230	240	250		
Dnmt3a2	201	caaggttaatt	gcagtaatga	atgctgtgga	agagaaccag	gcctctggag	250
DNMT3A2	201	caaggtcatt	gcaggaatga	atgctgtgga	agaaaaccag	gggccccggg	250
	260	270	280	290	300		
Dnmt3a2	251	agtctcagaa	ggtggaggag	gccagccctc	ctgctgtgca	gcagccacg	300
DNMT3A2	251	agtctcagaa	ggtggaggag	gccagccctc	ctgctgtgca	gcagccact	300
	310	320	330	340	350		
Dnmt3a2	301	gaccctgctt	ctccgactgt	ggccaccacc	cctgagccag	taggagggg	350
DNMT3A2	301	gaccggcat	ccccactgt	ggctaccacg	cctgagcccg	tgggttccga	350
	360	370	380	390	400		
Dnmt3a2	351	tgctggggac	aagaatgcta	ccaaagcgc	cgacgatgag	cctgagatag	400
DNMT3A2	351	tgctggggac	aagaatgcca	ccaaagcagg	cgatgacgag	ccagagatcg	400
	410	420	430	440	450		
Dnmt3a2	401	aggatggccg	gggctttggc	attggagagc	tggtgtggg	aaaacttcgg	450
DNMT3A2	401	aggacggccg	gggctttggc	attggggagc	tggtgtggg	aaaactgcgg	450
	460	470	480	490	500		
Dnmt3a2	451	ggcttctcct	ggtgccagg	ccgaatttgt	tcttgtgg	tgacagggccg	500
DNMT3A2	451	ggcttctcct	ggtgccagg	ccgcatttgt	tcttgtgg	tgacggggccg	500
	510	520	530	540	550		
Dnmt3a2	501	gagccgagca	gctgaaggca	ctcgctgggt	catgtggttc	ggagatggca	550
DNMT3A2	501	gagccgagca	gctgaaggca	cccgctgggt	catgtggttc	ggagacggca	550
	560	570	580	590	600		
Dnmt3a2	551	agttctcagt	ggtgtgtgt	gagaagctca	tgccgcttag	ctccttctgc	600
DNMT3A2	551	aattctcagt	ggtgtgtgtt	gagaagctga	tgccgcttag	ctcggtttgc	600
	610	620	630	640	650		
Dnmt3a2	601	agtgcattcc	accaggccac	ctacaacaag	cagcccatgt	accgcaaagc	650
DNMT3A2	601	agtgcgttcc	accaggccac	gtacaacaag	cagcccatgt	accgcaaagc	650
	660	670	680	690	700		
Dnmt3a2	651	catctacgaa	gtcctccagg	tggccagcag	ccgtgccggg	aagctgttcc	700
DNMT3A2	651	catctacgag	gtcctgcagg	tggccagcag	ccgcgcgggg	aagctgttcc	700

FIG. 13E-1

	710	720	730	740	750
Dnmt3a2	701	cagcttgc	tgacagtgtat	gaaagtgc	gtggcaaggc
DNMT3A2	701	cgttgtgc	cgacagcgat	gagagtgc	ctgccaaggc
	760	770	780	790	800
Dnmt3a2	751	cagaacaagc	agatgattga	atggggcc	ggtggcttcc
DNMT3A2	751	cagaacaagc	ccatgattga	atggggcc	gggggcttcc
	810	820	830	840	850
Dnmt3a2	801	tcctaagg	ctggagccac	cagaagaaga	gaagaatcc
DNMT3A2	801	ccctaagg	ctggagccac	cagaagaaga	gaagaatccc
	860	870	880	890	900
Dnmt3a2	851	tttacaccg	catgtgggt	gagcctg	cagctgc
DNMT3A2	851	tgtacacgg	catgtgggt	gaacctg	cagctgc
	910	920	930	940	950
Dnmt3a2	901	ccaccagcc	agaaaacc	aaagagc	acagagaa
DNMT3A2	901	ccaccagcc	aaaagcccc	gaagagc	gcggaga
	960	970	980	990	1000
Dnmt3a2	951	ggagatcatt	gatgagcg	caagggag	gctgggt
DNMT3A2	951	ggagattatt	gatgagcg	caagagag	gctgggt
	1010	1020	1030	1040	1050
Dnmt3a2	1001	agaagtgc	aaacatcg	gacatttgc	tctcatgt
DNMT3A2	1001	agaagtgc	gaacattgc	gacatctgc	tctcctgt
	1060	1070	1080	1090	1100
Dnmt3a2	1051	gtcaccc	agcaccc	cttcattg	ggcatgtg
DNMT3A2	1051	gttaccc	aacaccc	cttcgttgc	ggaatgtg
	1110	1120	1130	1140	1150
Dnmt3a2	1101	gaactgc	ttggagtgt	cttaccag	tgacgac
DNMT3A2	1101	gaactgc	ctggagtgt	cgtaccag	cgacgac
	1160	1170	1180	1190	1200
Dnmt3a2	1151	cctattgc	catctgc	ggggggc	aagtgc
DNMT3A2	1151	cctactgc	catctgc	ggggggc	aggtgc
	1210	1220	1230	1240	1250
Dnmt3a2	1201	aacaactgc	gcaggtgc	ttgtgtcg	tgtgtggat
DNMT3A2	1201	aacaactgc	gcaggtgc	ttgcgtgg	tgtgtggacc
	1260	1270	1280	1290	1300
Dnmt3a2	1251	gccaggagc	gctcaggc	ccattaagg	agacccctgg
DNMT3A2	1251	gccggggc	gcccaggc	ccattaagg	agacccctgg
	1310	1320	1330	1340	1350
Dnmt3a2	1301	tgtgcgggc	taagggcacc	tatgggc	tgcgaagac
DNMT3A2	1301	tgtgcgggc	caagggtacc	tacgggc	tgcggcggc
	1360	1370	1380	1390	1400
Dnmt3a2	1351	ccttctcgac	tccagatgtt	cttgccaat	aaccatgacc
DNMT3A2	1351	cctcccccgc	tccagatgtt	cttcgcta	aaccacgacc

FIG. 13E-2

	1410	1420	1430	1440	1450		
Dnmt3a2	1401	ccccccaaag	gtttaccac	ctgtccagc	tgagaagagg	aagcccatcc	1450
DNMT3A2	1401	ccctccaaag	gtttaccac	ctgtccagc	tgagaagagg	aagcccatcc	1450
	1460	1470	1480	1490	1500		
Dnmt3a2	1451	gcgtgctgtc	tctctttgat	gggattgcta	caggcctct	ggtgctgaag	1500
DNMT3A2	1451	gggtgctgtc	tctctttgat	ggaatcgcta	caggcctct	ggtgctgaag	1500
	1510	1520	1530	1540	1550		
Dnmt3a2	1501	gacctggca	tccaagtgg	ccgctacatt	gcctccgagg	tgtgtgagga	1550
DNMT3A2	1501	gacttggca	ttcaggtgg	ccgctacatt	gcctccgagg	tgtgtgagga	1550
	1560	1570	1580	1590	1600		
Dnmt3a2	1551	ctccatcacg	gtggcatgg	tgcggcacca	gggaaagatc	atgtacgtcg	1600
DNMT3A2	1551	ctccatcacg	gtggcatgg	tgcggcacca	gggaaagatc	atgtacgtcg	1600
	1610	1620	1630	1640	1650		
Dnmt3a2	1601	gggacgtccg	cagcgtcaca	cagaagcata	tccaggagtg	gggcccattc	1650
DNMT3A2	1601	gggacgtccg	cagcgtcaca	cagaagcata	tccaggagtg	gggcccattc	1650
	1660	1670	1680	1690	1700		
Dnmt3a2	1651	gacctggta	ttggaggcag	tccctgcaat	gacctctcca	ttgtcaaccc	1700
DNMT3A2	1651	gatctggta	ttggggcag	tccctgcaat	gacctctcca	tcgtcaaccc	1700
	1710	1720	1730	1740	1750		
Dnmt3a2	1701	tgcccgcaag	ggactttatg	agggtactgg	ccgcctcttc	tttgaggttct	1750
DNMT3A2	1701	tgctcgcaag	ggcctctacg	agggcactgg	ccgcctcttc	tttgaggttct	1750
	1760	1770	1780	1790	1800		
Dnmt3a2	1751	accgcctcct	gcatgatgcg	cggcccaagg	agggagatga	tcgcccattc	1800
DNMT3A2	1751	accgcctcct	gcatgatgcg	cggcccaagg	agggagatga	tcgcccattc	1800
	1810	1820	1830	1840	1850		
Dnmt3a2	1801	ttctggctct	ttgagaatgt	ggtggccatg	ggcgtagtg	acaagaggg	1850
DNMT3A2	1801	ttctggctct	ttgagaatgt	ggtggccatg	ggcgtagtg	acaagaggg	1850
	1860	1870	1880	1890	1900		
Dnmt3a2	1851	catctcgca	tttcttgagt	ctaaccctgt	gatgattgac	gccaaagaag	1900
DNMT3A2	1851	catctcgca	tttctcgagt	ccaaccctgt	gatgattgat	gccaaagaag	1900
	1910	1920	1930	1940	1950		
Dnmt3a2	1901	tgtctgctgc	acacagggcc	cgttacttct	gggtaaacct	tcctggcatg	1950
DNMT3A2	1901	tgtcagctgc	acacagggcc	cgctacttct	gggtaaacct	tcccggtatg	1950
	1960	1970	1980	1990	2000		
Dnmt3a2	1951	aacaggcctt	tggcatccac	tgtgaatgt	aagctggagc	tgcaagagt	2000
DNMT3A2	1951	aacaggccgt	tggcatccac	tgtgaatgt	aagctggagc	tgcaggagt	2000
	2010	2020	2030	2040	2050		
Dnmt3a2	2001	tctggagcac	ggcagaatag	ccaaagtccag	caaagtgagg	accattacca	2050
DNMT3A2	2001	tctggagcat	ggcaggatag	ccaaagtccag	caaagtgagg	accattacta	2050
	2060	2070	2080	2090	2100		
Dnmt3a2	2051	ccaggtcaaa	ctctataaaag	cagggcaaag	accagcattt	ccccgtcttc	2100
DNMT3A2	2051	cgaggtcaaa	ctccataaaag	cagggcaaag	accagcattt	tcctgtcttc	2100

FIG. 13E-3

	2110	2120	2130	2140	2150
Dnmt3a2	2101 atgaacgaga aggaggacat cctgtggtgc actgaaatgg aaagggtgtt	2150			
DNMT3A2	2101 atgaatgaga aagaggacat cttatggtgc actgaaatgg aaagggtatt	2150			
	2160	2170	2180	2190	2200
Dnmt3a2	2151 tggcttcccc gtccactaca cagacgtctc caacatgagc cgcttggcga	2200			
DNMT3A2	2151 tggtttccca gtccactata ctgacgtctc caacatgagc cgcttggcga	2200			
	2210	2220	2230	2240	2250
Dnmt3a2	2201 ggcagagact gctgggccga tcgtggagcg tgccggtcat ccgccacctc	2250			
DNMT3A2	2201 ggcagagact gctgggccgg tcatggagcg tgccagtcat ccgccacctc	2250			
	2260	2270	2280	2290	2300
Dnmt3a2	2251 ttgcgtccgc tgaaggaata ttttgcttgt gtgttaaggga catgggggca	2300			
DNMT3A2	2251 ttgcgtccgc tgaaggagta ttttgcgtgt gtgttaaggga catgggggca	2300			
	2310	2320	2330	2340	2350
Dnmt3a2	2301 aactgaagta gtgatgataa aaaagttaaa caaacaaaca aacaaaaaac	2350			
DNMT3A2	2301 aactgaggtt gcgac----a caaagttaaa caaacaaac- ---aaaaaac	2350			
	2360	2370	2380		
Dnmt3a2	2351 aaaacaaaac aataaaaacac caagaacgag				
DNMT3A2	2351 acaaaaacat- aataaaaacac caagaacatg				

FIG. 13E-4

Dnmt3a2	1	MNAVEENQASGESOKVEEASPPAVQOPTDPASPTVATTPEPVGGDAGDKN	50
DNMT3A2	1	MNAVEENQGPGESOKVEEASPPAVQOPTDPASPTVATTPEPVGSDAGDKN	50
Dnmt3a2	51	ATKAADDEPEYEDGRGFIGIGELVGKLRGFSWWPGRIVSWWMTGRSRAAE	100
DNMT3A2	51	ATKAGDDEPEYEDGRGFIGIGELVGKLRGFSWWPGRIVSWWMTGRSRAAE	100
Dnmt3a2	101	GTRWVMWFGDGKFSVVCVEKLMLPSSLFCASFHQATYNKQPMYRKAIYEVL	150
DNMT3A2	101	GTRWVMWFGDGKFSVVCVEKLMLPSSLFCASFHQATYNKQPMYRKAIYEVL	150
Dnmt3a2	151	QVASSRAGKLFPACHDSDESDSGKAVEVONKQMIEWALGGFOPSGPKGLE	200
DNMT3A2	151	QVASSRAGKLFPVCHDSDESDTAKAVEVONKPMIEWALGGFOPSGPKGLE	200
Dnmt3a2	201	PPEEEKNPYKEVYTDMWVEPEAAAYAPPPPAAKKPRKSTTEKPKVKEIIDE	250
DNMT3A2	201	PPEEEKNPYKEVYTDMWVEPEAAAYAPPPPAAKKPRKSTAEPKPKVKEIIDE	250
Dnmt3a2	251	RTRERLVYEVROKCRNIEDICISCGSLNVTLLEHPLFIGGMCONCKNCLE	300
DNMT3A2	251	RTRERLVYEVROKCRNIEDICISCGSLNVTLLEHPLFVGGMCONCKNCLE	300
Dnmt3a2	301	CAYQYDDDGYQSCTICCGGREVLMCGNNNCCRCFCVECDLLVPGAAQ	350
DNMT3A2	301	CAYQYDDDGYQSCTICCGGREVLMCGNNNCCRCFCVECDLLVPGAAQ	350
Dnmt3a2	351	AAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFANNHDQEFDPPKVY	400
DNMT3A2	351	AAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFANNHDQEFDPPKVY	400
Dnmt3a2	401	PPVPAEKRPPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVG	450
DNMT3A2	401	PPVPAEKRPPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVG	450
Dnmt3a2	451	MVRHQGKIMYVGDVRSVTQKHIQEWPFDLVIIGGSPCNDLSIVNPARKGL	500
DNMT3A2	451	MVRHQGKIMYVGDVRSVTQKHIQEWPFDLVIIGGSPCNDLSIVNPARKGL	500
Dnmt3a2	501	YEGTGRLFFFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFL	550
DNMT3A2	501	YEGTGRLFFFYRLLHDARPKEGDDRPF FWLFENVVAMGVSDKRDISRFL	550
Dnmt3a2	551	ESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGR	600
DNMT3A2	551	ESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGR	600
Dnmt3a2	601	IAKFSKVRTITTRSNSIKQGKDQHEPVFMNEKEDILWCTEMERVFGFPVH	650
DNMT3A2	601	IAKFSKVRTITTRSNSIKQGKDQHEPVFMNEKEDILWCTEMERVFGFPVH	650
Dnmt3a2	651	YTDXSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV	689
DNMT3A2	651	YTDXSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV	689

FIG. 13F

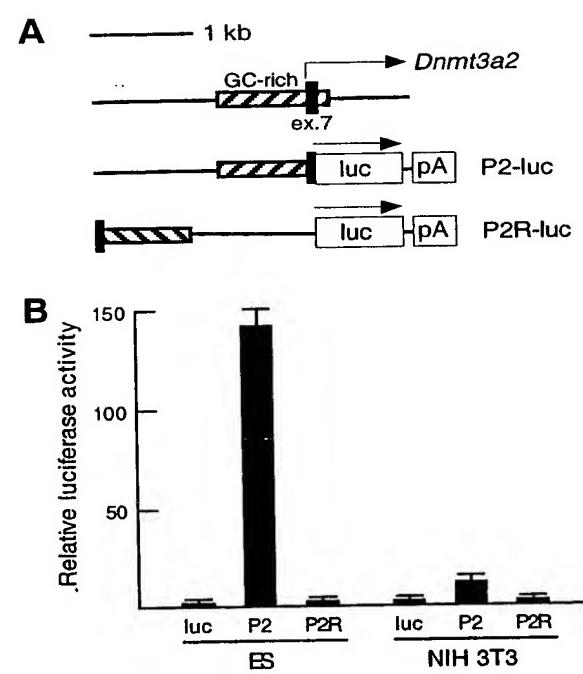


FIG. 14

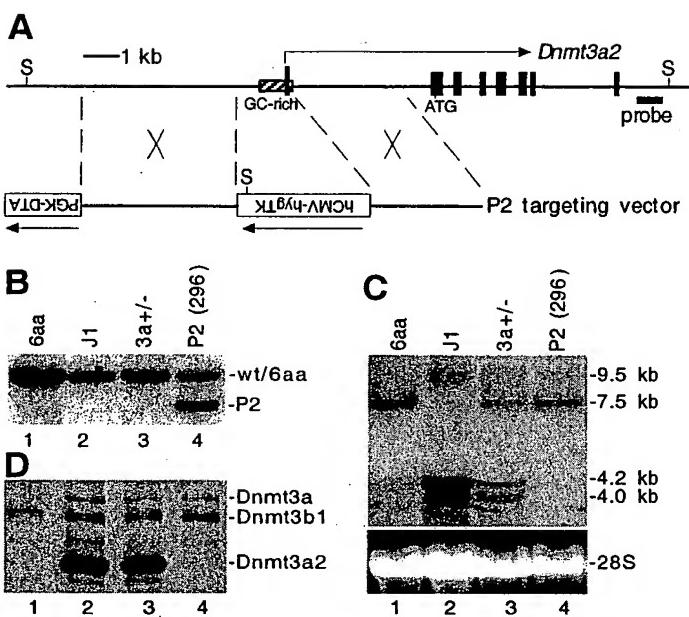


FIG. 15

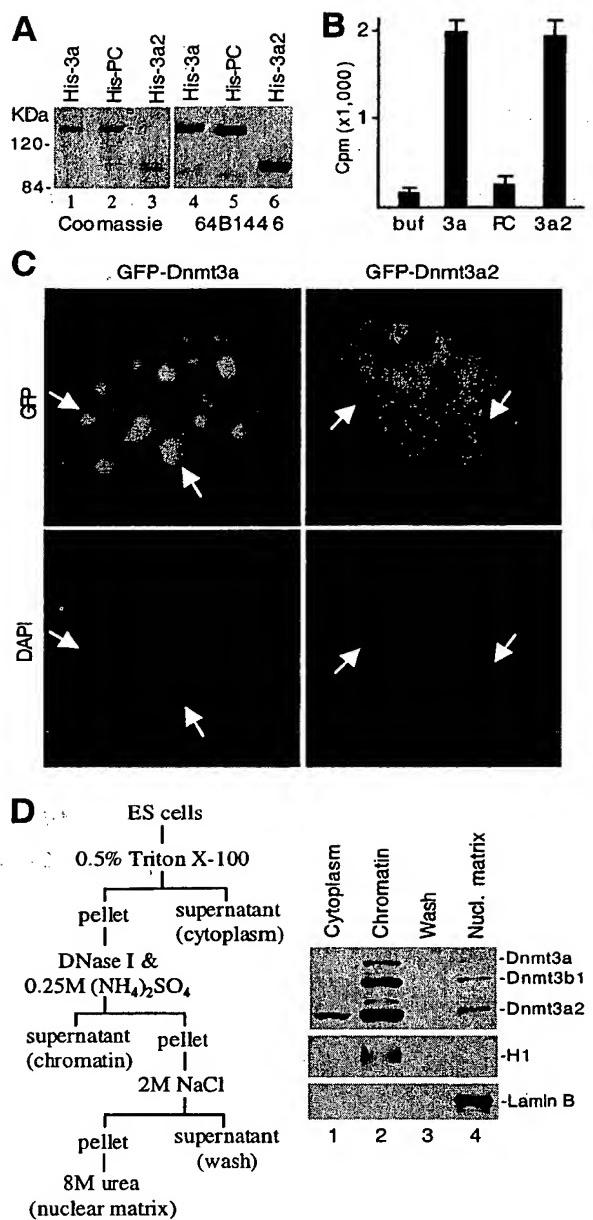


FIG. 16

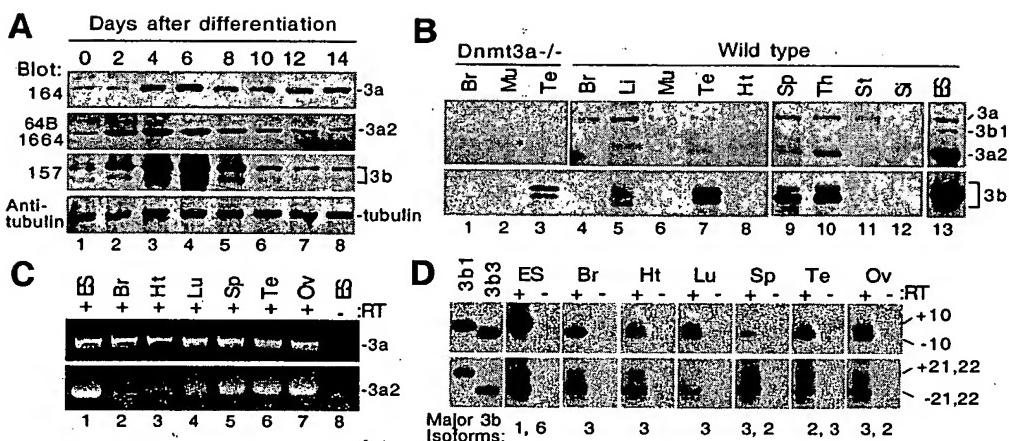


FIG. 17

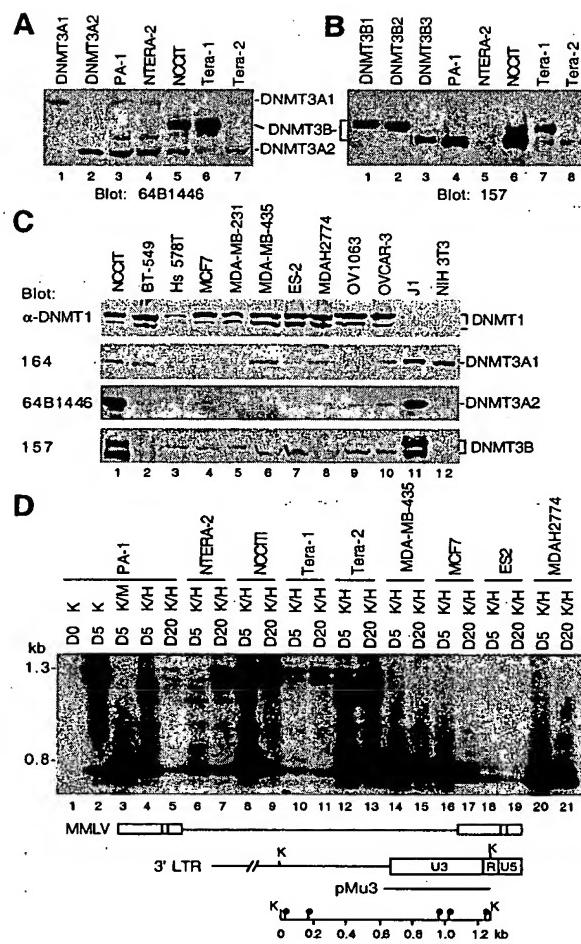


FIG. 18

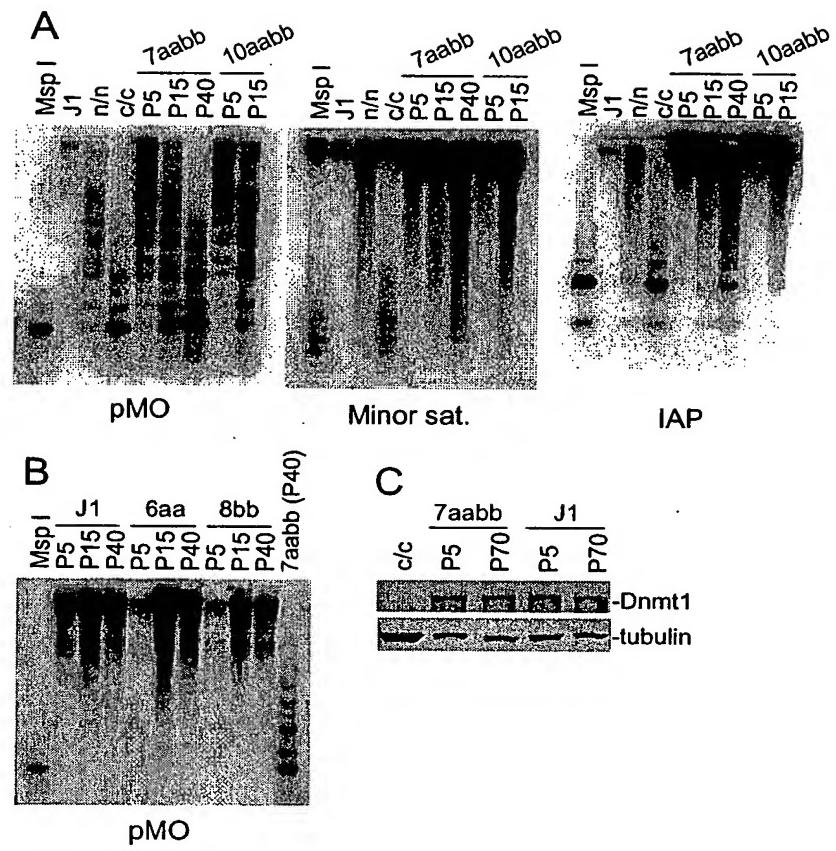


FIG. 19

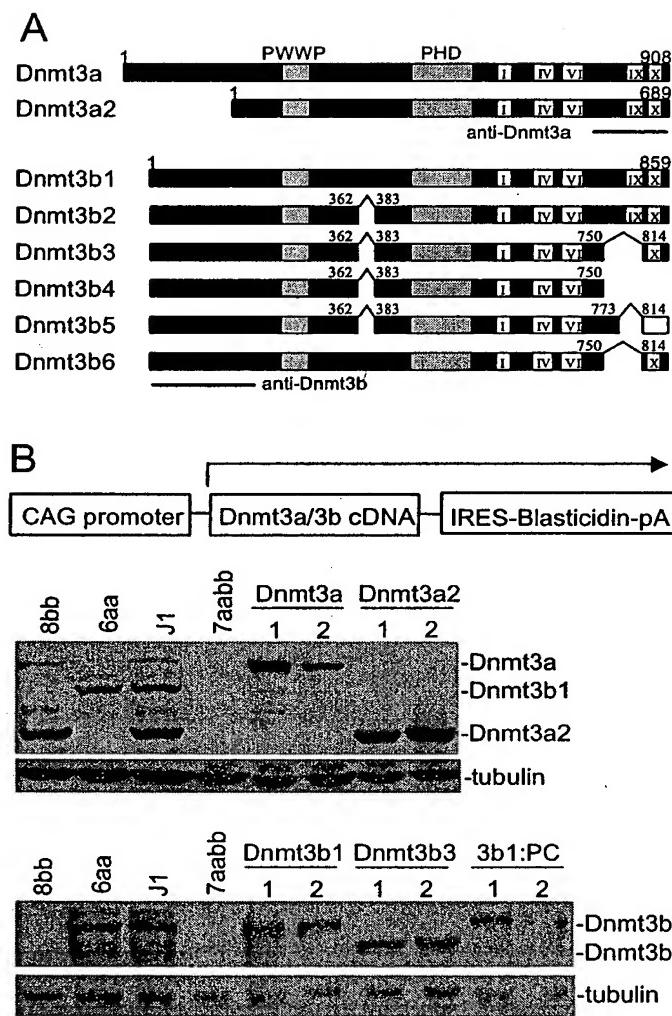


FIG. 20

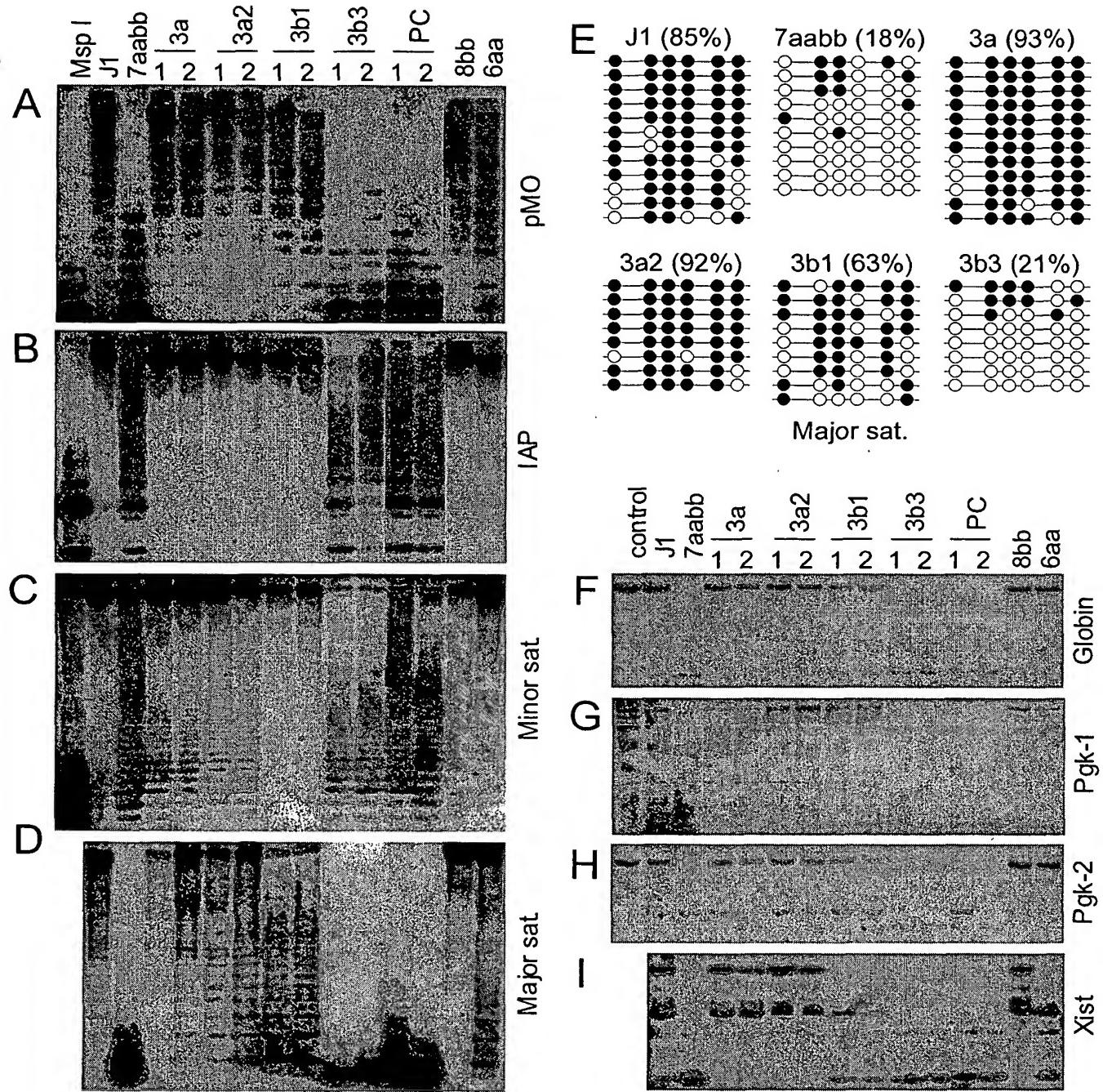


FIG. 21

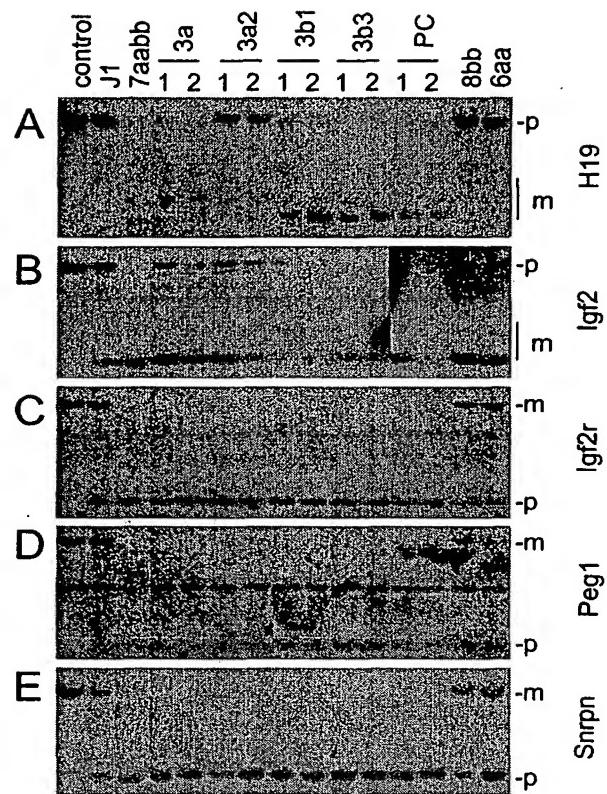


FIG. 22

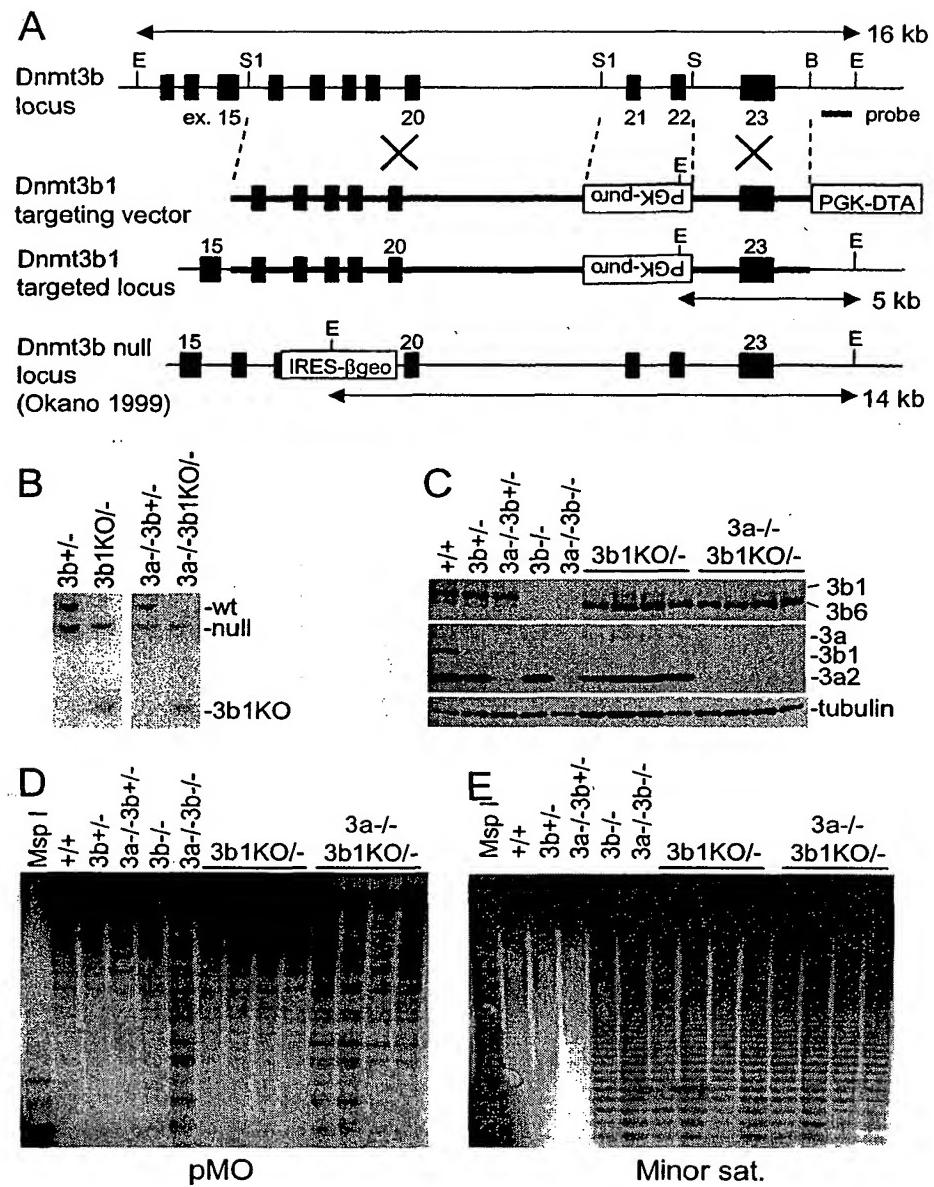


FIG. 23

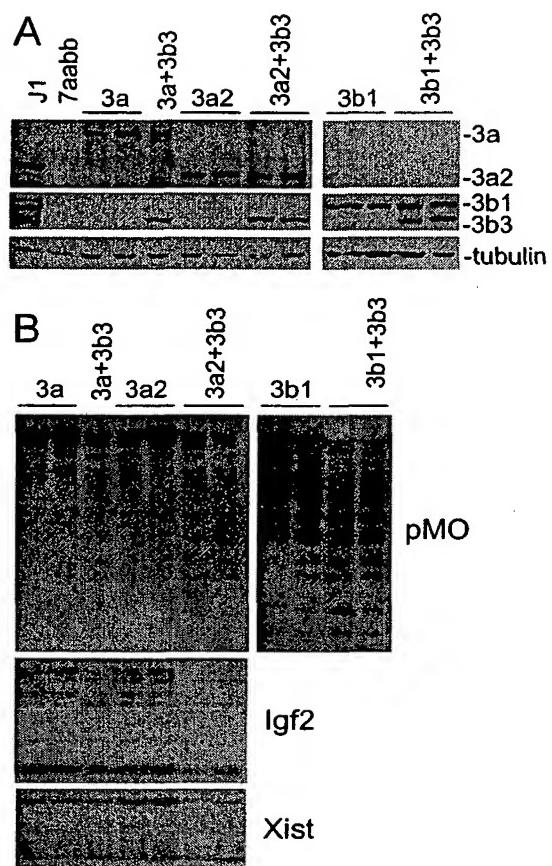


FIG. 24

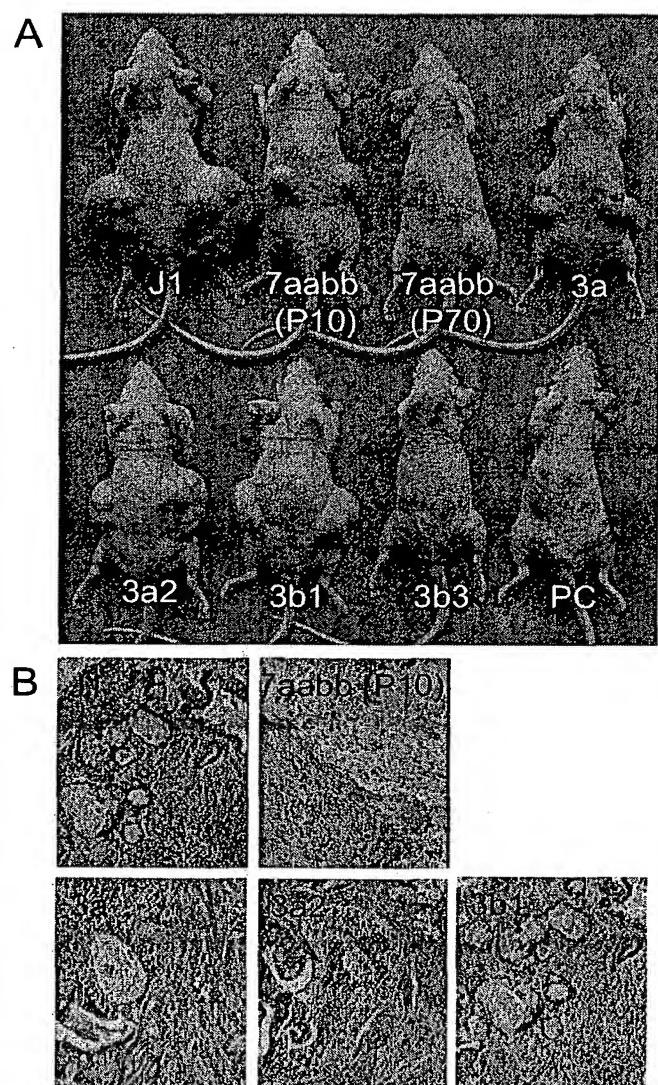


FIG. 25

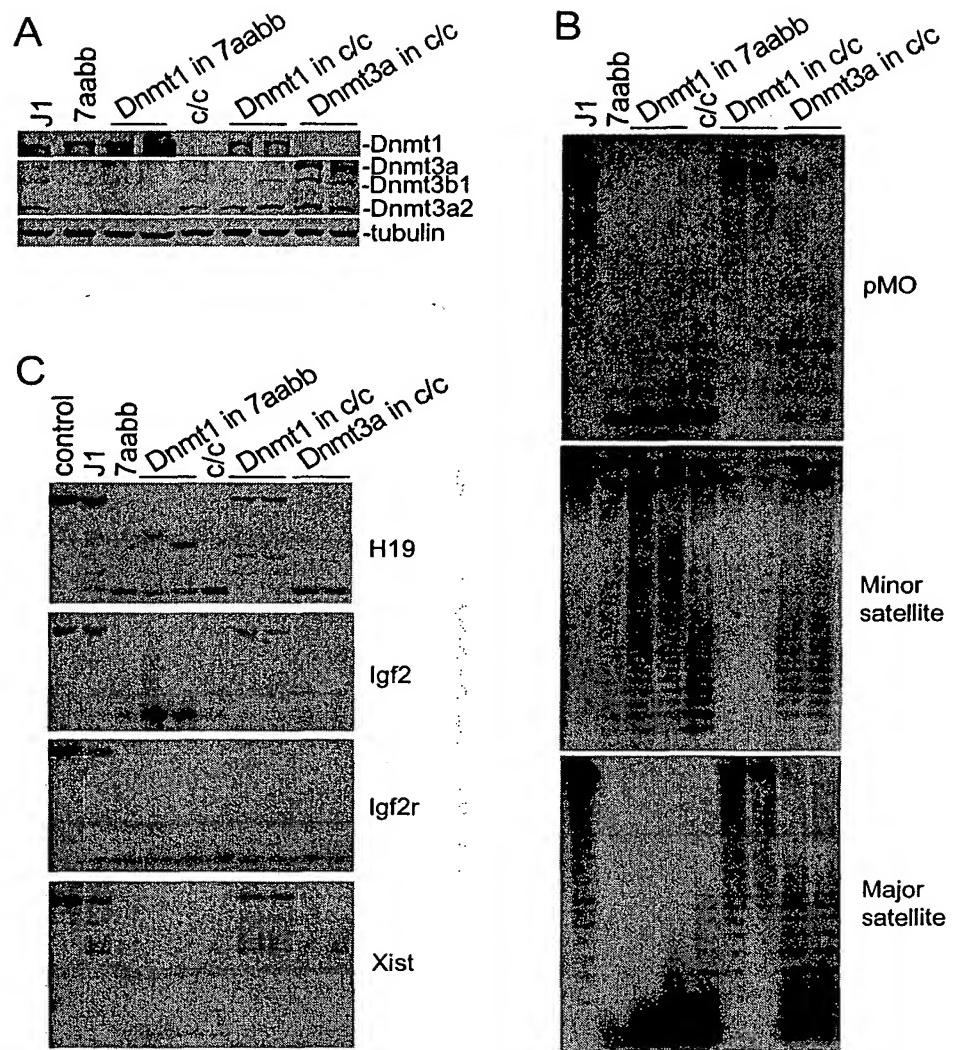


FIG. 26

Mouse Dnmt3a2 Promoter Sequence

1 GGAGCCAGGCACCTGGGTGTTACCTCAGTGCCTTAGGATATTGGTTTCAGCTCT
60 AGAGGGCTGATGTCATCACCCCTATTTGAGATGAGAAAACAGACATCTTGGGTTAA
119 GTGGTCTGTGTCAGGTCACCGCAATGGGATCAGGTCTTCCCAAGCGTCCAGCCAGA
178 TAGCGCGGCTCCCTGCTGGGCATTCTCCTCAGTTCTTGTCTAATTCATCTTGCA
237 AACTTAATCCTGGCTAACATCTTGAAAATACTCATTACCTTGTTCCAGAACATCT
296 GCCATGTTACAGAATATCTCCATTCACTGCTTGACCCCAGTCCCACACTCAGCCATT
355 AGCTTAGTAAAATTGAGAGGGTGGGTGGAAGAGTTCTTCTCCTTACCTGCTTG
414 CCACCTCAAATCGTGGTTATCTCTGATCTACTGTCCTATCTCTCACCCACACCCT
473 TCATTGATGCAGCCTCTGCTATCTGCTTGGTGGTTGGTAGTTATCCACACAGGAG
532 TTTGCTTTCACTGATTCCCCCTCCCCCACCCATCTCCCAAGTCTAGTGGAAATCTA
591 TCAACCTCCTGAGAGCAGGACCAAGTGTCCATTCTGTATCCGATGATGCTCCAGTCCT
650 CTAATGGGGGGGGGGCGGGCGCCAGGAGTGGCGTGTGCTTCTCAAAACCCAACCT
709 TAGTCCTCTACTGTNN
768 NNN
827 NNN
886 NNN
945 NNNNNNNNNNNNNNNNNNNNNNNNNNNACACACCTGGGTACTATGTCTTGCTCAG
1004 GAATGGCATGAAATGGCCTACACTTACCTGGTGGTTCTAGGAGAGAGACACTAGCAG
1063 TGCGTGGAGTGTCTATTACTATTACATAATTGCTGAGACAGGGTTCTGATGTTCT
1122 AGGCTGGCCTTGAACTTGTGTTAGTCAGAACATGATCTTAAATTCTGATCTGGTTCT
1181 CCAAGTCTAGGATTACAGGTGTACTTCACCACCAAAAGTTGAACAGCTGCAGATGCC
1240 TTGGCATTGCTCTAACGAACAGAAAATGAAACAAGCAAGCAAGACCCATTGTGACCCG
1299 GGGGACTCGGGGACTGGACGGGAAGTTCAAAGTCTACTTGTGAACCACGCTTTTA
1358 AAGCACCCCCCTCCATTACCTGTAGCGTGGCGGTGAAGTTATTGTCTGGGCGCCCTC
1417 AACCTGCGTGGGACACCTCTATCCACTCACATCTGCTTGTACTTGCCTAAACTAC
1476 GTTCCGTAAACTCCGAGCCTCATCTCTAACATCTGAAACTTGCTAGCGCGCTCGCAC
1535 GCGCTTTTTTTTTTCCGGAAACTCACTTCTACAAACTTCTCCCGGAC
1594 TCTCAGGCTGTCAGGCCAGCGCTCCTGTCCTCCACCACCGCTGCTCTGGTGCCCCGCG
1653 GCCCGCACGCACCCCTGCCTCCCTCAAGGTCCCCAACTTCCCTATGTACCCCCCCCAC
1712 CAGAGTTGGGGAAAGGGAGCAGAGCGGGCTGTCCCTAAACCTGGCTGGAGGGGGGG
1771 CCCTGGGAACGGACTGGCCAGCCTCTCCCCCAGGGCCCCCGGCCCTGGGCCGGGT
1830 GAGGGCTGGCCCAGCGCCAGCGTAGGAGGCCGGCCCTCCCCCCGGCCCGCTTAG
1889 CCAACCAGAAACTCCAGTGGGGCCACGTGACCTGGAGTTCTAGACAAAGAAAATGTTCT
1948 CCTCCCTCCCCCGGGCCCCCTCCCCCTCCCTCTGGCCCCCTCCGCCCCCAACCCCA
2007 ACGCCCTGCCCCCTCCCCCAGACGGCAGCTATTACAGAGCTTCGGGCCGGGCTC
2066 ACACCTGAGCTGTA
CTGCAGAGGGCTGCACCTGGCCTTATGG

FIG. 27

Human DNMT3A2 promoter sequence

1 GGAGCCAGGCACCTAGAGAATTGTCTCATTGTCATTAGGAGATGGTGGCGTCCATG
60 GCCAAAGAGGGCTGATGTCATCACTCGTTGCAGATGAGACAACAGATTTCTGGG
119 GGTAAAGTGACTTGTAAAGGTATGGTGGAAACAGAACTGAAGTCCAGATCTT
178 TTTTTTTTTTTGAGACGGAGTCTCGCTCTGTTGCCAGGCTGGAGTCAGTG
237 GCATGATCTCGCTCACTGCAACATCCGCCTCTAAAGTTGAAGCGATTCTCTGCCT
296 CAGCCTCCCAGTAGCTGGGATTACTGGCGCACGCCACCACGCCTGGCTAATTTTG
355 TATTTTAGTAAAGACAAGGTTCACCATGTTAGTCAGGCCGTCTCAAACCTCTGA
414 CCTCATGATCCGCCTGCTCAGCCTCAAAGTGTGGGATTATAGGCGTGAGCCAC
473 CGCGCTCGGCCAAGTCCAGATCTCTAACAAAGTGGCGTGCCTAAAGTCCCTCTGC
532 TGTGGGGTGCATTTCCATTTCCTCAGTTCTCCTCTAATTCATCTTGCCAAC
591 GGCAACTAGGCTGATTTCAAAACTCATTATCTTGTCAGAAAACCTGCGGTT
650 ATTCTCCCTGCTACAGAATATAACCAAGGACGCACCTGAAGGCTTGCCATTACCTT
709 GCCCTGTCGTGACTGGGAGGGTGGAGGTGGCGAGGGTCTCCTCCCTCCCCAGCCC
768 GGCAGCTCTGCTCATCCTACCCATCTCACCTCATTCAAAGTCCGATCCAGCCTCCA
827 GGGCCAGTCGGCTCACCTGGAAGTGCACCTCTGACCTCTTGTATCCATGCCGCC
886 ATTTTTTCTACTTGGTATTGTGGCATAGTTACCTTACATATGTTGTTTACAG
945 TGATCCTTCATATTCTCAAGTCTAGTGGAACTTCACACCCCTCGAGGGCAGAGC
1004 CAACAGGGCTATTTCTATCTGATCCTACAGCCAACGTAATGGAGGGCTGTTGGT
1063 GGGGACTGCGTCTGCCTTGGGGTAGGTGCCTTGTTCAGGAGGGAGAGCTTGAAA
1122 TGGCGGAGGCTGCACCTGGAGGCCGACCTGGAGGCCAGGAGAGGAGTCAGGTCT
1181 TCTCGATCTGCAGATGTTGAGCCTGGGAATGAAGGAATTGCTGAACCTTCTGAAGG
1240 AGCGCCCTCGCCGCGACCAACCTGCAAACAGGAAATGAGAAATCCAGGGAAAGGCC
1299 CAGAGTGACGCAGGGGCCCTGGGACTCGAACGCTGACCTCTCACGCCGCGCTTTT
1358 GAGGCCCCCCCGCTCTCTATTCACCTGTAGTGTGGAGGCCAGGCCAAACA
1417 ATCCCCGATCTGGAGCGCTCCCAATGCCTGCGCGCCTGCTGTCACTCTCCGTCTG
1476 TGTGCTGAGTTTCTACAGCTCCTGGCCTCCTATCTGTAAGCTTTCTTTT
1535 TTTTTTGGTTGTGCTTCAGAGAAACTCACTTTACAACCTTCTCCGGCTCTCCC
1594 AGGCCGTCCGAAAGCTCGGCTTGCTTCCGGACCCCCGGCTCCCTCCGGCAG
1653 GCGGCTCGGGAGCAGCCCTCCCTCCCGCCGGCCCCGGCGCTTAATCT
1712 CTTCCAGAGCTGGGGAGGGCCAGGCGGTCTCCCGAAGGCAGGGGCGCTCCCTGCA
1771 GCCCGGCCTGGCGGGCCCTGGGAACGGCGGGGAACGGCCTCGCCCCCGGCCCG
1830 CGCCCTCGGACCGGAGAAGAGGGCTGGCCCAGCGCCAGCGTCGGAGGCCGGCC
1889 CCTCCCCGGGCCGCTCGCAGCCAACCAGGCCCTCCAGCGGGGCCACGTGACCTGGA
1948 GTCCTAGACAAAGAAAATGTTCCCTCCCTCCCCCGCCGCCCTCCCTCCAG
2009 TGGCCCCCTCCGCCCCCAGCCCCATCGCCCCCTTCCCTCCCCAAGACGGGCAGCT
2066 ACTTCCAGAGCTTCAGGGCCGCGCTCACACCTGAGCGCAGTCAGAGGGCTGCA
2125 CCTGGCCTTATGG

FIG. 28

Mouse and human Dnmt3a2 promoter alignment

Top Sequence = mouse Dnmt3a2 promoter, 1858 bp (gap not counted)
Bottom Sequence = human DNMT3A2 promoter, 2065 bp

1-104 (1-105) 77% ==
 1289-1338 (1475-1530) 82% ==
 1518-1858 (1724-2065) 87%

FIG. 29